WHITEPAPER

Untargeted Feature Extraction Optimized for Breath Analysis

Authors: Madeleine Ball, Hannah Winter, Daniel Tuck, Holly Whittome, Ibrahim Karaman
Owlstone Medical, Cambridge, UK

Keywords: Breath Biopsy, GC-MS, untargeted metabolomics, feature extraction, volatile organic compounds, breath biomarkers

Key Points

- Owlstone Medical has developed an optimized untargeted feature extraction workflow for the OMNI® platform
- This workflow has an enhanced ability to handle retention time shifts seen during larger or longer clinical studies
- Assessment of molecular features has been automated to increase efficiency and allow for more selective manual review
- Metabolite identification has been aligned with the MSI guidelines for reporting the untargeted feature tables

Introduction

Exhaled breath is a promising, but currently underutilized sampling medium in clinical research and practice. Some biomarkers that can be analyzed using breath are already in clinical use, including hydrogen, methane, and nitric oxide (1), and there have been hundreds of clinical trials that have reported promising results associating compounds in breath with clinically relevant biological pathways. Breath samples are inherently complex and contain a large number of volatile compounds that originate both from being inhaled from the environmental air, and from metabolic processes within the body. Because this is a fundamental characteristic of breath, robust analytical approaches are needed that can cope with this complexity. Once breath samples have been collected and stored in sorbent tubes during our studies, they are processed at Owlstone Medical through thermal desorption-gas chromatography-mass spectrometry (TD-GC-MS) on high-resolution accurate mass (HRAM) Q Exactive Orbitrap systems. This generates a GC-MS dataset, and characteristic patterns of peaks can be used to measure and identify compounds through cross-referencing with our in-house Breath Biopsy VOC Atlas, as well as generic NIST (The National Institute of Standards and Technology) online libraries. This workflow has been successfully used to advance the development of several compounds in breath that have great potential to be utilized as biomarkers, such as limonene as a biomarker for liver function (2,3). However, there are limitations to the current analytical workflow that could be improved upon, including reducing the need for manual review, which could maximize the speed and effectiveness of the analysis. A proprietary, optimized untargeted feature extraction workflow has now been developed internally at Owlstone Medical to address these issues and will be discussed in detail in this whitepaper.

Read the Breath Biopsy OMNI Whitepaper

Owlstone Medical developed Breath Biopsy OMNI® (Owlstone Medical Novel Insights) assay, an end-to-end pipeline that allows the robust identification and measurement of the volatile organic compounds (VOCs) present in exhaled breath (summarized in Figure 1).

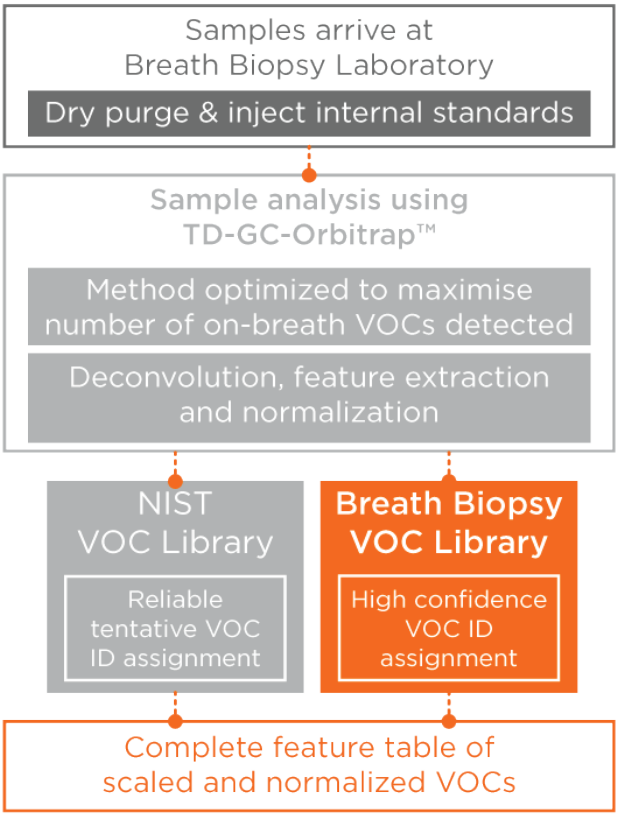

Samples arrive at Breath Biopsy Laboratory
Dry purge & inject internal standards
Sample analysis using TD-GC-Orbitrap™
Method optimized to maximise number of on-breath VOCs detected
Deconvolution, feature extraction and normalization
NIST VOC Library
Reliable tentative VOC ID assignment
Breath Biopsy VOC Library
High confidence VOC ID assignment
Complete feature table of scaled and normalized VOCs

Figure 1: The processing pipeline of breath samples as part of our untargeted OMNI® service.

Data processing that encompasses statistical analysis, identification, and interpretation of potential biomarkers in the breath is a crucial part of Owlstone Medical's OMNI® workflow. As well as targeted analysis for specific compounds, the nature of discovery work on breath often calls for much broader untargeted analyses that can identify unexpected changes in compound concentrations, while maximizing the number of compounds that can be detected. This untargeted analysis of breath VOCs is performed by deconvolving and extracting unknown chemical compounds from GC-MS spectra, enabling subsequent assigning of compound identifications (IDs) to molecular features (MFs) in the dataset, along with their relative concentrations.

General untargeted feature extraction for breath analysis involves the following challenges:

- Retention time misalignments may occur when the peak shift is high across the samples and can be a source of bias in poorly balanced datasets
- Manual review is time-consuming and subjective
- Quality metrics for peak detection (peak shape, peak symmetry) are not always available
- Background correction is not always possible or reliable
- Gap-filling processes in the data are not always optimal

The objective of the Owlstone Untargeted Feature Extraction project was to develop an optimized, fast, and scalable untargeted feature extraction method that could be used in all future OMNI® studies. Although the previous method provided satisfactory outcomes, manually reviewing each extracted feature can affect the delivery time and cost. Therefore, we aimed to develop an untargeted feature extraction method that decreases the delivery time and cost by automating the manual review process. While achieving this, our purpose was further to improve the processing steps to extract high-quality features from raw TD-GC-MS data acquired from breath samples.

The new feature extraction workflow aimed to have the following attributes:

- Accurate and reproducible peak/compound detection
- Repeatability across samples, including where retention time alignment is necessary
- Minimized false negatives and positives.
- Ability to assign tentative IDs with high confidence
- Faster, more efficient delivery
- Ability to handle a high number of samples
- Availability of recording, auditing, and visualization

Results

An Automated Feature Extraction Solution

The new untargeted feature extraction workflow involves typical processing steps such as centroiding, baseline correction, peak deconvolution and peak grouping, retention time alignment, gap filling and compound identification. We added extra feature validity filters and improvements to the retention time alignment and gap filling steps. Furthermore, we aimed to align the features IDs with MSI standards.

The new untargeted feature extraction workflow can be applied on a GC-MS dataset and ends with the generation of a feature table with the peak areas of the quantifier ions, deconvoluted mass spectra and IDs with tiers of confidence (Figure 2).

After peaks have been generated and grouped into compounds, a validation step must be undertaken to quality control the data before compounds can be identified. Furthermore, the features should have correspondence across the samples. This step was previously undertaken manually, which was time-consuming and required specialist analytical personnel.

The steps for the new workflow were defined as follows:

1. Validity criteria for the deconvoluted peaks

The first step of the new workflow was to generate

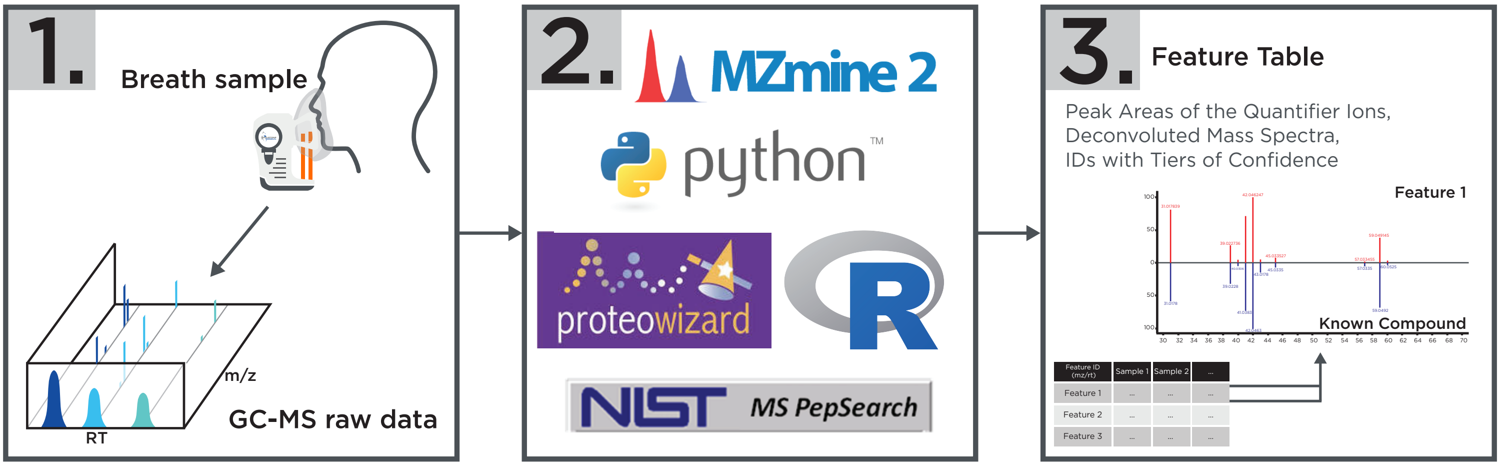


Figure 2: An overview of the Owlstone Medical Untargeted Feature Extraction Workflow.

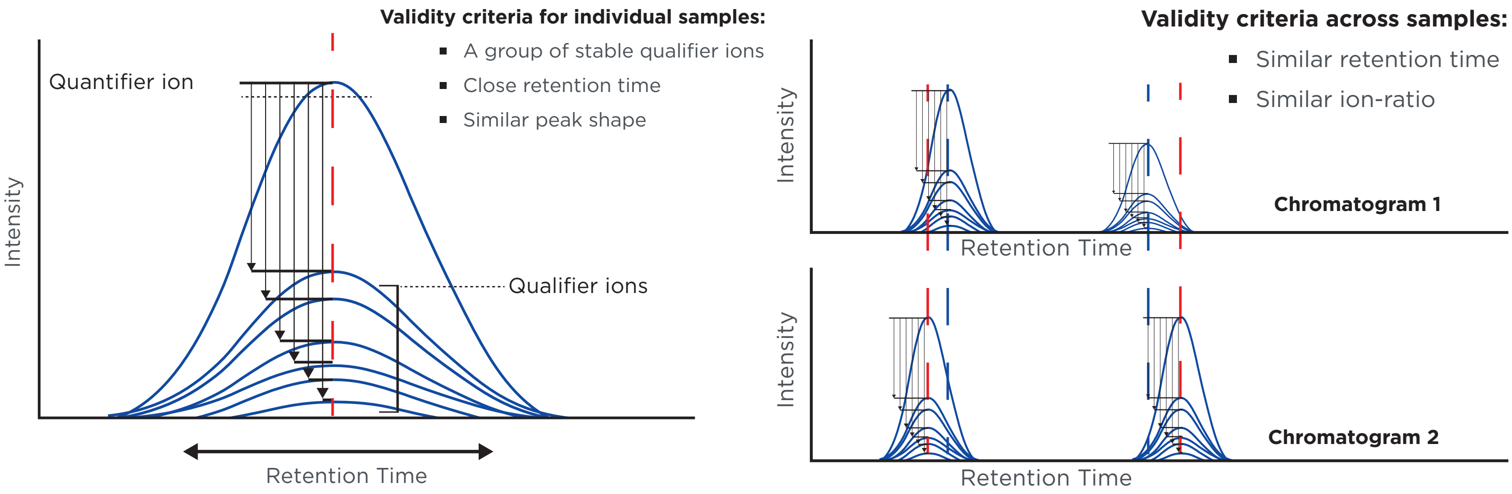


Figure 3: An overview of the analytical properties of validity criteria in an individual sample, and across other samples.

peaks and group them into compounds using optimized parameters. Additionally, we defined validity criteria that can be applied automatically to filter the deconvoluted peaks that generate the features (Figure 3). In other words, features of the dataset were first detected, and from these, validity criteria were developed that can be applied both for each individual sample and across samples.

Typical criteria for individual samples are:

- A group of stable qualifier ion peaks (with signal-to-noise-ratio above a threshold)
- Close retention time (within a retention time range)
- Similar peak shape (full-width-at-half-maximum values of qualifier ions within three standard deviations of the mean, or equivalent such as three median absolute deviations to the median)

In addition to those above, the criteria across the samples are:

- Similar retention time (retention times across the samples within three standard deviations of the mean, or equivalent)
- Similar ion-ratio (within an ion-ratio range)

These validity characteristics aim to target and filter out artefactual features within the samples. Furthermore, they will help uncover the retention time and fragmentation patterns across the samples and determine the outlying samples that occurred to be included in the feature table due to the untargeted nature of retention time alignment.

The sample validity criteria can then be applied as a validity filter to each sample prior to retention time alignment. Subsequently, the validity criteria across the samples can be used on the features after the feature table is generated to remove features that do not meet the requirements. This was previously undertaken manually, and so an automated process in the Owlstone Untargeted Feature Extraction Workflow to perform these validity checks in a systematic manner can improve and speed up the processing of GC-MS data.

2. Improvements in merging duplicated or split features

After valid features are detected for every sample, retention time alignment is performed to have the same compounds in a feature across the samples. It is common to observe duplicated or split features after retention time alignment because the alignment algorithm looks for features with similar mass spectra to align into a single feature. Slight changes in mass spectra due to inconsistent fragmentation patterns across the samples or co-eluted compounds in the samples might cause duplicated or split features. To tackle this type of issue in the data, the Owlstone Untargeted Feature Extraction Workflow searches for features with the same m/z in a predefined retention time interval and aims to find common ions with similar ion-ratios across the samples of those features. A new feature (or a set of features) is then generated where duplicated or split features are merged into one main feature and others if there are remaining ones that cannot be merged.

3. A novel gap-filling approach

After all the preprocessing steps, the resulting feature table may still contain gaps due to poor peak detection or misalignment. The fact that there are missing sample peaks in a feature does not prove that the peak does not exist. For a certain sample that has a gap in a feature, the gap-filling process involves searching a peak within a retention time and an m/z window. This approach does not assure assigning a peak that corresponds to a valid feature. Therefore, in the Owlstone Untargeted Feature Extraction Workflow, gap-filling is achieved by searching among the already detected peaks and alternatively among the deconvoluted spectra that represent VOCs. Further, the peaks or deconvoluted spectra found are subjected to validity checks. This approach results in gap-filled cells with features consisting of 3 levels defined by confidence/reliability:

– Level 1: Reference mass found and at least 2 qualifier ions, which have similar ion-ratios.

– Level 2: Reference mass found and at least

2 qualifier ions, but these do not have similar ion-ratios.

– Level 3: Only the reference mass is found.

If a reference mass is found among the compounds with deconvoluted spectra, the gap-filled values are tagged as "Level 1A" and "Level 2A". On the other hand, if a reference mass and at least 2 qualifier ions are found among the raw peaks within an RT window, the gap-filled values are tagged as "Level 1B" and "Level 2B". Finally, if only the reference mass can be found among the raw peaks, the gap-filled values are tagged as "Level 3".

Note that the conventional gap-filling methods for LC-MS and GC-MS data provide only Level 3 gap-filling.

4. VOC ID assignment

Compound identification is challenging to standardize, and so is often referred to as the bottleneck of metabolomics. A standard was first outlined in 2007 by Sumner et al (4), which provided a top-level framework for different levels of confidence for identification. Different interpretations of the standards outlined have subsequently been published (5–7), and we have internally analyzed the performance of many of these.

To ensure the highest quality VOC identification, Owlstone Medical has taken a stringent interpretation of the MSI standard (8) – the Metabolomics Standard Initiative carried out by an international community of volunteers to create broad community consensus. This results in assigned IDs given in terms of four broad tiers of confidence (Figure 4).

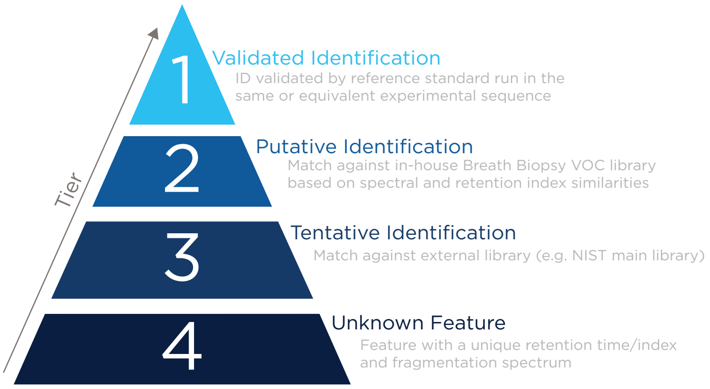

Tier
1
Validated Identification
ID validated by reference standard run in the same or equivalent experimental sequence
2
Putative Identification
Match against in-house Breath Biopsy VOC library based on spectral and retention index similarities
3
Tentative Identification
Match against external library (e.g. NIST main library)
4
Unknown Feature
Feature with a unique retention time/index and fragmentation spectrum

Figure 4: An overview of the different tiers of confidence that are assigned to identified compounds.

The highest confidence ID assignment is Tier 1. For a compound to be categorized into Tier 1, there must have been a reference standard run in the same sequence as the breath sample to ensure identical experimental conditions. Chromatographic (retention time) matching of the sample to the standards improves the confidence in addition to the high-resolution mass spectral matching.

In the absence of reference standards, the application of our Breath Biopsy VOC Atlas produces Tier 2 IDs, known as "putative identification". An assignment of a compound with Tier 2 is less confident than Tier 1 but still reliable as putative IDs due to the advantages of high-resolution mass spectral matching.

Tier 3 are tentative IDs that are assigned based on cross-referencing to the NIST library containing over 40,000 compounds. Mass spectral matching is applied in low resolution to the data and tentative IDs are assigned.

Validation of the Untargeted Feature Extraction Workflow

The new workflow for untargeted feature extraction was tested both for accurate peak/ compound detection, and accuracy across samples.

1. Validation of accurate compound detection

To evaluate the performance of the Owlstone Untargeted Feature Extraction Workflow for accurate peak/ compound detection, a set of breath samples spiked with a standard mix containing 121 compounds plus 8 internal standard compounds were utilized (Figure 5). In addition, the samples were spiked with 5 different levels of concentrations using the standard mix. The gap-filled feature table and library match table were generated by the workflow for this sample set and the features corresponding to the spiked compounds were found by comparing each feature against the Owlstone HRAM library. For each spiked compound (not including internal standard compounds) an R^2 value was generated between the peak area values and the concentration spiked on the sample.

Validation of accurate compound detection

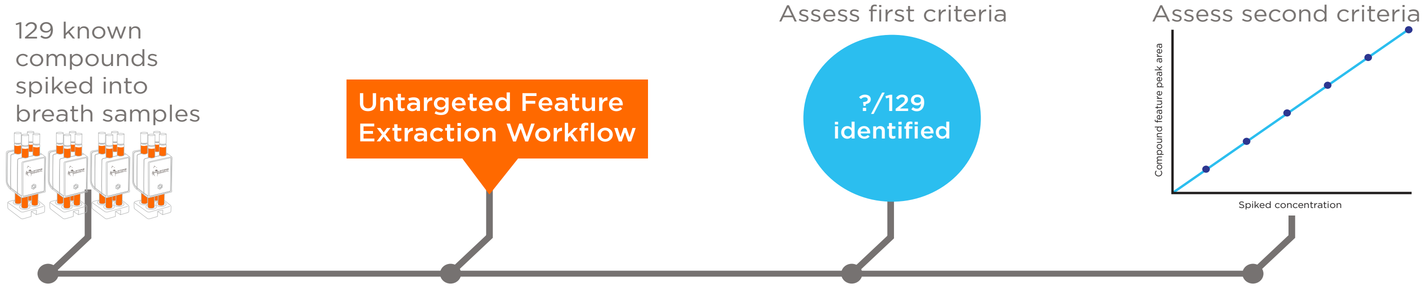

129 known compounds spiked into breath samples
Untargeted Feature Extraction Workflow
Assess first criteria
?/129 identified
Assess second criteria
Compound feature peak area
Spiked concentration

Figure 5: An overview of the validation study to test the accuracy of the untargeted feature extraction workflow to detect spiked compounds.

Two criteria were considered necessary to pass validation:

1. **At least 95% of the compounds spiked onto the sample can be identified.**
2. **Over 90% of those compounds found have a peak area that is linearly correlated with the concentration on the tube.**

There were 129 compounds spiked in the Feature Extraction development samples, of these 124 were found in the gap-filled feature table, representing 96.1% of compounds. The compounds that were not found were mainly due to limitations in the chromatography, i.e., co-eluting compounds causing suboptimal peak deconvolution.

Of the 124 compounds found, 8 were internal standards with constant concentration and 116 were spiked in three levels of concentration (in addition to a non-spiked level). The compounds which were not internal standards were plotted to see the correlation between the peak area values from the gap-filled feature table and the concentration which was spiked onto tubes. A total of 99 compounds have an R^2 >= 0.95, which is 85.3% of the compounds found. The reason for not having a good correlation for several compounds was the relatively high variation in baseline concentrations between the subjects compared to the spiked concentrations, e.g., limonene. In addition, compounds such as acetic acid producing tailing or fronting peaks were found to interfere with the compounds eluting nearby.

Based on the results, keeping the limitations of the untargeted nature of the analytical method in mind, **both validation criteria were deemed to be successful. It is challenging to accurately measure each compound in complex samples such as breath by a single untargeted analytical method. Nonetheless, the accuracy can be enhanced by improving the samplers by including a filter for specific compounds such as acetic acid or by further optimizing the analytical method.**

2. Validation of accuracy across samples

To test the Owlstone Untargeted Feature Extraction Workflow for accuracy across samples, a set of breath samples from an internal study were used (Figure 6). This study was a 5 weeklong study that included breath samples from 4 volunteers. In total 9 breath samples and 9 equipment blank samples (using the ReCIVA® mask and CASPER® airflow) were collected per volunteer. The samples were spiked with 8 internal standards. Targeted data (Chromeleon output) generated for 52 standard compounds and 8 internal standards compounds was considered as reference data. This dataset and the output of Owlstone Untargeted Feature Extraction Workflow for the target compounds are expected to be either the same or highly correlated.

To test the peak integration accuracy of the Owlstone Untargeted Feature Extraction Workflow across the samples, the following criterion was considered to pass validation:

1. **At least 90% of the tested compound molecular features have an R^2 value of >0.95 when targeted Chromeleon data and Owlstone Untargeted Feature Extraction Workflow outputs are compared.**

By comparing the data of standard tubes, the criteria of 90% with an R^2 of 0.95 or above has been met with 95% of standards having an R^2 above 0.95. There were only three failing compounds , but all still showed a strong positive correlation with R^2 values above 0.92 with no outliers. Based on the results, the acceptance criteria were met using the reference standards data.

Conclusion

Breath samples are complex and produce feature-rich chromatograms. The untargeted feature extraction workflow we have developed provides a fast, versatile, scalable, and validated tool that can be used in future OMNI® breath analysis studies in an automated manner.

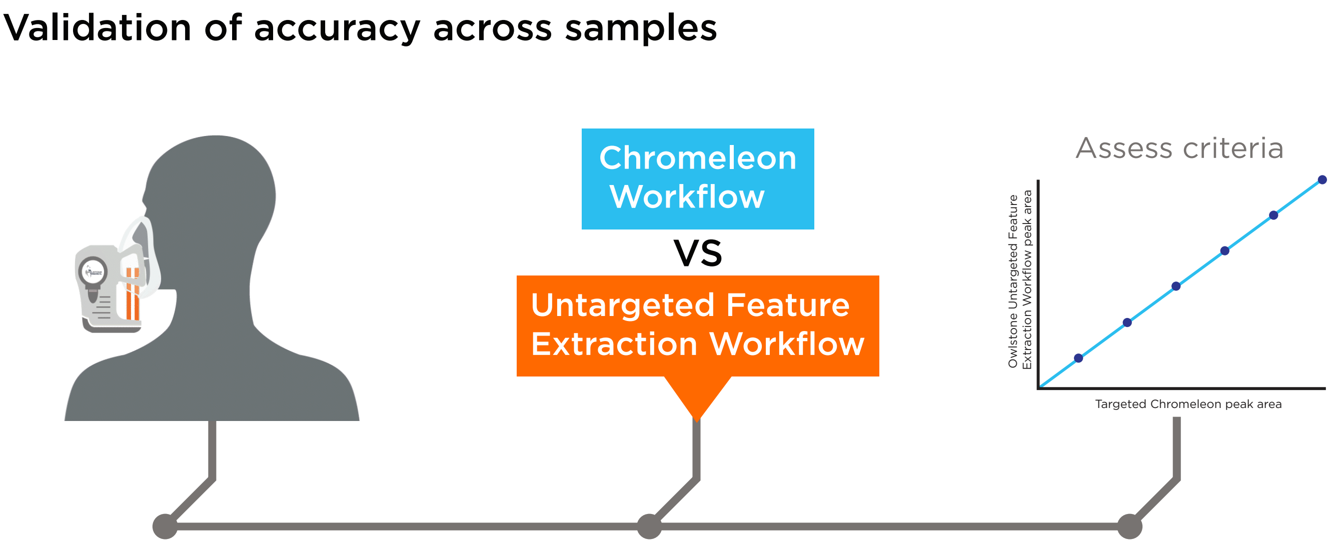

Validation of accuracy across samples
Chromeleon Workflow
VS
Untargeted Feature Extraction Workflow
Assess criteria
Owlstone Untargeted Feature Extraction Workflow peak area
Targeted Chromeleon peak area

Figure 6: An overview of the validation study to test the accuracy of the untargeted feature extraction workflow across samples.

References

1. Kiss H, Örlős Z, Gellért Á, Megyesfalvi Z, Mikáczó A, Sárközi A, et al. Exhaled Biomarkers for Point-of-Care Diagnosis: Recent Advances and New Challenges in Breathomics. Micromachines. 2023 Feb;14(2):391.

2. Fernández del Río R, O'Hara ME, Holt A, Pemberton P, Shah T, Whitehouse T, et al. Volatile Biomarkers in Breath Associated With Liver Cirrhosis — Comparisons of Pre- and Post-liver Transplant Breath Samples. EBioMedicine. 2015 Jul 26;2(9):1243–50.

3. Ferrandino G, Orf I, Smith R, Calcagno M, Thind AK, Debiram-Beecham I, et al. Breath Biopsy Assessment of Liver Disease Using an Exogenous Volatile Organic Compound—Toward Improved Detection of Liver Impairment. Clinical and Translational Gastroenterology. 2020 Sep;11(9):e00239.

4. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis. Metabolomics. 2007 Sep 1;3(3):211–21.

5. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. J Am Soc Mass Spectrom. 2016 Dec;27(12):1897–905.

6. Rochat B. Proposed Confidence Scale and ID Score in the Identification of Known-Unknown Compounds Using High Resolution MS Data. J Am Soc Mass Spectrom. 2017 Apr 1;28(4):709–23.

7. Ford L, Mitchell M, Wulff J, Evans A, Kennedy A, Elsea S, et al. Clinical metabolomics for inborn errors of metabolism. Adv Clin Chem. 2022;107:79–138.

8. Fiehn O, Robertson D, Griffin J, van der Werf M, Nikolau B, Morrison N, et al. The metabolomics standards initiative (MSI). Metabolomics. 2007 Sep 1;3(3):175–8.

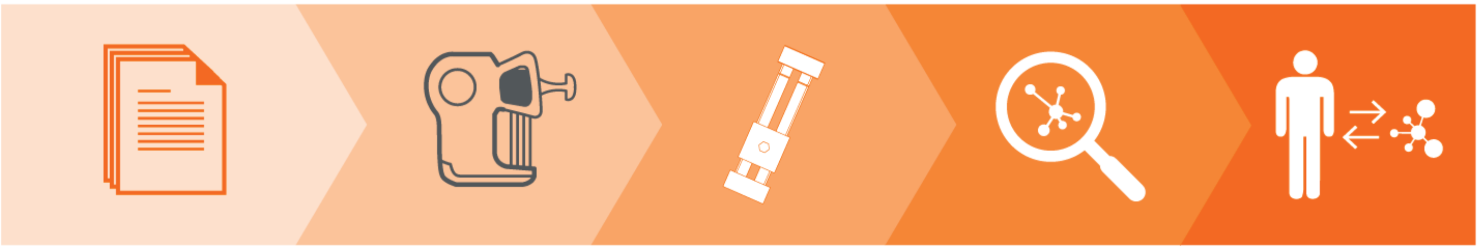
Add Breath Biopsy OMNI to your research

OMNI is available worldwide for clinical trials and academic research applications. If you are interested in finding non-invasive biomarkers for applications including early detection, precision medicine or drug development, get in touch to discuss studies with our specialist team.

To get started, scan the QR code to visit the OMNI webpage and find out more about its features and capabilities.

BREATH BIOPSY®
OMNI

The complete end-to-end breath VOC analysis service



Expert Study Design & Management

Robust Breath Collection

Reliable Sample Processing & Analysis

In-depth Data Analysis

Specialist Data Interpretation

Contact us to discuss identifying and validating breath biomarkers, and how to incorporate Breath Biopsy in your research.

breathbiopsy@owlstone.co.uk

OWLSTONE MEDICAL

Owlstone Medical Ltd, 183 Cambridge Science Park,
Milton Road, Cambridge, CB4 0GJ, UK

@owlstonemedical

Company Number 04955647 | VAT Number 260449214

Owlstone Medical's Products and Services are for research use only. Not for use in diagnostic procedures.

BREATH® BIOPSY