**Tackling the analytical challenges to establish a routine workflow for breathomics research**

Thibault Massenet1, Judith Potjewijd2, Rachid Tobal2, Fanny Gester3, Delphine Zanella1, Monique Henket3, Makon-Sébastien Njock3, Thibaut Dejong1, Gregory Gridelet1, Laurie Giltay3, Françoise Guissard3, Béatrice André4, Clio Ribbens4, Renaud Louis3, Pieter Van Paassen2, Jean-François Focant1, Julien Guiot3\*, Pierre-Hugues Stefanuto1\*

1 Molecular System, Organic & Biological Analytical Chemistry Group, University of Liege, Belgium.

2 Department of Internal Medicine, Division of Clinical and Experimental Immunology, Maastricht University Medical Center, Maastricht, The Netherlands.

3 Respiratory Medicine, CHU Liege, Belgium.

4 Rheumatology department, CHU Liège, Belgium.

\*co-last authors

Breathomics, an emerging field, investigates the intricate relationship between volatile organic compounds (VOCs) contained in exhaled breath and human health, offering a non-invasive approach to disease monitoring and detection. Analytical chemistry stands as a crucial component in the comprehensive characterization of the complex molecular composition of exhaled breath. However, challenges related to sampling or quality control still need to be addressed to establish routine clinical protocols. In this context, several community initiatives aim to establish guidelines for exhaled breath analysis[1]. A typical breathomics workflow includes critical steps such as sampling protocols, characterization methods, and data processing.

In this study, we present a comprehensive analytical workflow utilizing bag-based sampling combined with thermal desorption, coupled with comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry (TD-GCxGC-HRTOFMS). We developed sampling kits and standard operating procedures (SOPs) to reduce sampling variability. Additionally, we established common GCxGC-HRMS conditions compatible with all breathomics studies and worked on SOPs for data processing for this type of dataset.

Once the workflow fully developed, we conducted a multicentric clinical study to test the robustness of our approach, focusing on patients with systemic sclerosis (SSc) and investigating the potential of exhaled breath to monitor the development of interstitial lung disease (ILD), a major clinical complication in SSc. Forty-two patients (21 SSc, 21 SSc-ILD) were prospectively recruited from University Hospital of Liège (CHU), Belgium, and Maastricht University Medical Center (MUMC+), the Netherlands. Sampling kits were sent to both centers, and analyses were performed using a single instrumentation.

This study demonstrated that ready-to-use kits has helped to reduce cross-contamination, while reducing the workload of care staff and thus potential errors. This process also increased the number of patients sampled in a defined period reducing variability caused by longitudinal study. TD-GCxGC-HRTOFMS enabled the detection of around 700 features, with high inter-day repeatability confirmed using a mixture containing 21 standards covering a wide range of the chromatograph. Each of these standards was carefully selected to mimic exhaled breath composition based on previous studies[2]. This confirmed the identification of features but also allowed us to proceed to instrumental maintenance or intensities correction when needed. Finally, a breath-based model, based on nine markers, was developed. This model achieved an AUC (Area Under the Curve) of 0.82, accuracy of 85%, sensitivity of 77% and a specificity of 100% for identifying ILD phenotype. Among them, four markers were also previously highlighted in our previous study[3]. In addition, correlation between functional respiratory parameters and the VOCs confirmed the model’s ability to classify patients accurately. QC and cross-validation were also conducted to exclude overfitting problems or data structure modification. Random feature selection and class assignment permutations were performed to rule out overfitting.

In conclusion, this work represents a major step in demonstrating the potential of breathomics in a clinical setting. The implementation of SOPs in medical centers located in two different countries facilitated a comparative analysis of breath from patients with SSc and SSc-ILD. The discovery of potential markers contributes to a better understanding of the disease and metabolic pathways involved, offering the prospect of rapid targeted treatments for ILD patients. The multicentric design provided valuable insights into the robustness of our analytical methods, contributing to the development of guidelines and SOPs for larger exhaled breath studies. This marks a significant stride towards integrating breathomics into clinical settings.

**Bibliography**

[1] B. Henderson *et al.*, “A benchmarking protocol for breath analysis: the peppermint experiment,” *J. Breath Res.*, vol. 14, no. 4, p. 046008, Jul. 2020, doi: 10.1088/1752-7163/ABA130.

[2] F. N. Schleich *et al.*, “Exhaled volatile organic compounds are able to discriminate between neutrophilic and eosinophilic asthma,” *Am. J. Respir. Crit. Care Med.*, vol. 200, no. 4, pp. 444–453, Aug. 2019, doi: 10.1164/RCCM.201811-2210OC/SUPPL\_FILE/DISCLOSURES.PDF.

[3] D. Zanella *et al.*, “Breathomics to diagnose systemic sclerosis using thermal desorption and comprehensive two-dimensional gas chromatography high-resolution time-of-flight mass spectrometry,” *Anal. Bioanal. Chem.*, vol. 413, no. 14, pp. 3813–3822, Jun. 2021, doi: 10.1007/S00216-021-03333-4/FIGURES/3.