Real-time breath metabolomics to assess early response to CFTR modulators in adults with cystic fibrosis: an open-label proof-of-concept study

Emmanuelle Bardin,[1,2,3,4,*] Hélène Salvator,[3,5,6,*] Camille Roquencourt,[2,3,*] Elodie Lamy,[2] Nicolas Hunzinger,[2] Isabelle Sermet-Gaudelus,[1,4] Sandra De Miranda,[6] Dominique Grenet,[6] Philippe Devillier,[3,5] Stanislas Grassin-Delyle[2,3]

[1]Université Paris Cité, INSERM U1151, CNRS UMR8253, Institut Necker Enfants Malades, Paris, France; [2]Université Paris-Saclay, UVSQ, INSERM, Infection et inflammation (2I), U1173, Département de Biotechnologie de la Santé, Montigny le Bretonneux, France; [3]Hôpital Foch, Exhalomics®, Suresnes, France; [4]Hôpital Necker Enfants Malades, Centre de Référence Maladies Rares, Paris, France; [5]Université Paris-Saclay, UVSQ, UFR Simone Veil-Santé, VIM-Suresnes, UMR0892, Suresnes, France; [6]Hôpital Foch, Service de pneumologie, CRCM - Centre de Transplantation Pulmonaire, Suresnes, France.

* co-first authors

Corresponding author: Stanislas Grassin-Delyle; Exhalomics, Hôpital Foch; 40, rue Worth; 92150 Suresnes; France; E-mail address: s.grassindelyle@hopital-foch.com; Phone: +33.1.46.25.73.93.

Word count: abstract: 249 words; text: 3435 words

SUMMARY

Real-time PTR-MS breath analysis revealed rapid metabolic changes in adults with cystic fibrosis starting ETI therapy. Several VOCs shifted towards healthy levels within one month and correlated with lung function and sweat chloride. These findings support breathomics as a non-invasive surrogate for early monitoring of CFTR modulator response.

ABSTRACT

Background and objective: Highly effective CFTR modulators, particularly elexacaftor/tezacaftor/ivacaftor (ETI), produce rapid clinical improvements in people with cystic fibrosis. Yet early treatment effects may be difficult to capture with spirometry or a sweat test in patients with a mild disease or atypical mutations. Exhaled breath is rich in volatile organic compounds (VOCs) reflecting metabolic and inflammatory processes. We aimed to determine whether ETI induces early, measurable changes in breath composition and whether these changes relate to clinical outcomes. **Methods:** Ten adults initiating ETI were enrolled in a prospective, open-label study with breath sampling at baseline, week one and month one. VOCs were measured using real-time proton-transfer-reaction – mass spectrometry (PTR-MS). Longitudinal changes were assessed using multilevel statistics, including univariate linear mixed-effects models and multivariate RM-ASCA+; repeated-measures correlations examined associations with lung function and sweat chloride concentration. Results were compared with a healthy cohort. **Results:** Amongst the eight responders, 11 VOC features changed significantly after ETI initiation. Eight differed from healthy controls at baseline and shifted towards healthy levels over one month. RM-ASCA+ identified

linear and non-linear temporal patterns capturing acute and progressive metabolic responses. A 14-feature PLS-DA model classified visits with high accuracy (AUC=0.96–1.00). Ten VOCs correlated with clinical readouts. Features of interest were tentatively identified and pointed towards a shift in the microbiome and/or energy metabolism. **Conclusion:** ETI induces rapid alterations in exhaled VOCs, many trending towards healthy values and correlating with clinical improvement. Real-time breath analysis offers a promising non-invasive surrogate for early monitoring of therapeutic response.

Trial registration: NCT05295524

Keywords: Breathomics; cystic fibrosis; CFTR modulators; real-time mass spectrometry; volatile organic compounds.

INTRODUCTION

The development of cystic fibrosis transmembrane conductance regulator (CFTR) modulators is transforming the therapeutic landscape for patients with cystic fibrosis (CF). In eligible patients, the combination of the potentiator ivacaftor with the two correctors tezacaftor and elexacaftor (ETI) achieves unprecedented improvements in lung function and quality of life within a few weeks [1]. Correctors support the intra-cellular processing of misfolded CFTR proteins, which result from class II mutations such as the most common p.Phe508Del (F508del thereafter), whilst the potentiator improves the functioning of CFTR channels positioned at the apical side of epithelial cells. In pulmonary cells, the restoration of CFTR activity reinstates the ionic osmotic balance, which translates into rehydration of lung-lining fluids and restoration of mucociliary clearance [2]. ETI is indicated for patients with at least one non-Class I mutation that produces a CFTR protein amenable to correction.

However, these highly efficient therapies came with poorly anticipated challenges. Clinical evaluation of people with CF (pwCF) mainly focuses on respiratory function through the monitoring of the forced expiratory volume in one second (FEV_1) and the forced vital capacity (FVC), since the respiratory status still mostly determines prognosis. Yet, these endpoints sometimes fail to detect clinical benefit or biological changes, for example in pwCF with high baseline lung function and in young children unable to perform these tests or, conversely, in people with irreversible lung damage [3,4]. The measurement of sweat chloride concentration (SCC) often serves as a secondary endpoint in clinical trials as a marker of CFTR activity, although it does not always correlate with clinical improvement, e.g. in respiratory function and symptoms [5,6]. Indeed, some patients show delayed or limited effects on lung function or SCC

[5,7], making these outcomes unreliable for assessing drug efficacy in certain cases. Nevertheless, other clinical benefits are emerging, such as reduced exacerbation frequency or lower bacterial load [8], normalisation of inflammation [9], or improved nutritional status [10], although long-term data are still lacking on the newest CFTR modulators combinations. Therefore, there is a clear rationale for developing alternative methods capable of capturing early therapeutic effects of novel treatments such as ETI across patients with diverse mutations and varying disease severity.

Over the past fifty years, breathomics has expanded across multiple domains of respiratory medicine and has emerged as a promising, innovative analytical approach. Exhaled breath constitutes a unique, individual-specific biological signature, rich in personalised metabolic information, sensitive to subtle physiological changes, and obtainable entirely non-invasively [11]. Exhaled metabolites, commonly referred to as volatile organic compounds (VOCs), are believed to arise from inflammatory processes, oxidative reactions, and enzyme-driven metabolic pathways, all of which may be modulated by pharmacological treatments. These compounds have been associated with a wide range of conditions, including chronic and acute pathologies, infections, ageing, environmental exposures, and therapeutic interventions [12]. Notably, a previous study following up 20 F508del-homozygous patients initiating the CFTR modulator dual therapy lumacaftor/ivacaftor reported significant metabolic changes in exhaled breath profiles after three months which persisted at 12 months [13]. Online proton-transfer reaction – mass spectrometry (PTR-MS) is a high-sensitivity, point-of-care technology allowing real-time monitoring of VOCs which has already demonstrated potential for the diagnosis of infectious or cancerous diseases [14–16]. We hypothesised that ETI-

induced metabolic modifications could be detected in exhaled breath using PTR-MS, and that alterations in volatilomic profiles might be associated with clinical outcomes.

METHODS

Study design

We conducted a prospective, open-label study at the Foch university hospital (Suresnes, France), enrolling any adult with CF (≥18 years) initiating ETI. Patients were participating to the PHEAL-KAFTRIO study, approved by an ethics committee (Comité de protection des personnes Sud-Est I, 2021-A03119-32, NCT05295524). All patients provided written informed consent. The study included three visits: before ETI initiation (V0), during the first week (V1) and after one month (V2) of treatment. At each visit, clinical readouts including respiratory symptoms (cough, sputum production, dyspnoea) and episodes of exacerbation were recorded. Spirometry outcomes (FEV_1 and FVC in absolute and percentage of predictive values (pp)) were measured according to American Thoracic Society /European Respiratory Society guidelines [17]. Sweat collection was performed using a Macroduct® sweat collection system according to the recommendations of the European Cystic Fibrosis Society – Clinical Trial Network [18]. Response to ETI was defined by functional criteria: +10% $ppFEV_1$ and/or a drop of more than 20 mmol/L in SCC at V2. Patients with compassionate access to ETI, i.e. carrying mutations excluded from the initial indication, were evaluated after two months by an adjudication committee [5].

Breath analysis

At each visit, breath volatile metabolites were analysed with real-time mass spectrometry using a proton-transfer reaction – quadrupole interface – time-of-flight mass spectrometer (PTR-Qi-TOF MS; Ionicon, Innsbruck, Austria). Patients blew directly into the instrument via a single-use mouthpiece connected to a BET-med device

maintained at 70°C to avoid condensation. A hollow cathode ion source produces reagent ions (H_3O^+) and protons are transferred from H_3O^+ to VOCs in a drift tube forming $[VOC+H]^+$ ions based on proton affinity. A quadrupole then conveys and focuses the protonated analytes towards the TOF analyser, which separates them according to their mass-over-charge ratio (*m/z*). Operating parameters were as follows: electric field strength (E/N) 127.8 Td; source voltage 120 V; drift tube pressure, temperature and voltage were set to 3.8 mbar, 60°C and 959 V, respectively. Mass spectra were acquired up to *m/z* 392 with a time resolution of 1 s, experimental mass resolution was around 5000 [15]. Patients performed five exhalations of approximately 15 s with 40 s intervals, resulting in a quick procedure (<5min) requiring minimal patient cooperation.

The data from this cohort were compared with those obtained using the same instrument in a cohort of healthy young adults (VOC-COMPARE study, NCT06020521, 40 individuals enrolled between July-August 2023).

Breath analysis data processing

The data was processed using the ptairMS R package, following a previously reported workflow [15,19] including mass calibration, exhalation phase detection, peak deconvolution and quantification, peak alignment, background filtering, and missing value imputation. Quantitative values are reported as counts per second (cps) after probabilistic quotient normalisation (PQN) [20,21], log-transformation and autoscaling. Isotopes and saturated ion signals were removed; groups of correlated *m/z* features corresponding to water clusters or dehydrated molecules were identified ($\Delta m/z$=18). Putative annotations were proposed based on the human breathomics database [22] and

supported by others' published work [23–25] and an in-house VOC library. More details are provided in Appendix S1 and Figure S1 in the Supporting Information.

Statistical analysis

The effect of treatment was studied on the subgroup of responders only whilst correlations were assessed using data collected from all patients. The full statistical workflow is shown in Figure S2 in the Supporting Information.

To assess whether there was a systematic effect of the treatment on breath VOCs, we first performed a longitudinal univariate analysis using a linear mixed-effect model (LMM) for each feature, with the time as a fixed effect (visits as categorical variable), and patients as a random effect. An F-test (lmerTest R package [26] using Satterthwaite's method) was applied to the fixed-effect coefficients to test whether the evolutions between V0 and V1 or between V0 and V2 were significant.

Then, a repeated measures ANOVA–simultaneous component analysis plus (RM-ASCA+) was performed [27,28], consisting of a principal component analysis (PCA) applied to the LMM fixed-effect coefficients of each feature. This approach enables visualisation of the main temporal evolution patterns in the data, and quantification of the proportion of variance explained by time by averaging the marginal R_2 of the LMM models. More details are provided in Appendix S1 in the Supporting Information.

In addition, a PCA was performed to identify the main source of variation and a multilevel partial least-squares – discriminant analysis (PLS-DA) model was applied using the mixOmics R package [29,30], with visits as the predictive features and individuals as the multilevel factor. The model's performance was evaluated through 5-fold cross validation repeated three times, along with the computation of the area under

the curve (AUC) of the receiver operating characteristic (ROC) for each visit. One model was built using all features, a second on selected features, including the 10 features with the highest variable importance in projection (VIP) scores and the variables showing significant time effects according to the LMM (p-value <0.05).

A Wilcoxon-Mann-Whitney test was implemented to compare the expression of features at V0 and V2 to their expression in the cohort of healthy young adults. Correlations with clinical readouts ($ppFEV_1$, ppFVC, SCC) were assessed using a repeated measure correlation test (rmcorr [31]). All p-values from univariate tests were adjusted for multiple testing using the Benjamini–Hochberg (FDR) correction.

RESULTS

Study population and clinical response

Ten participants (two females), median age 34 years (interquartile range (IQR) 17), were enrolled from March to December 2022; three had compassionate access and two switched from monotherapy ivacaftor (Table 1). At baseline, lung disease stage was variable; all patients were positive for at least one respiratory pathogen (*Pseudomonas aeruginosa* (5/10), *Staphylococcus aureus* (5/10)). Seven had daily anti-infectious prophylaxis (inhaled and oral route) throughout the study (Table S1 in the Supporting Information). Over the course of the study, three patients experienced a mild bronchial exacerbation; two of them were related to a coronavirus disease (COVID-19) infection and required additional antibiotics (Table S1 in the Supporting Information).

At V2, six patients carrying F508del experienced clinical improvement and met clinical response criteria (Table 1). One patient switching from ivacaftor chose to stop ETI after one month because of insomnia; despite a +8% gain in $ppFEV_1$ his response was deemed inconclusive. Amongst the three patients who were granted off-label ETI access, one exhibited a quick response, the second showed a delayed response with +10% $ppFEV_1$ after two months, and the third discontinued the treatment based on poor clinical outcomes. This resulted in a total number of 8/10 responders. Side effects were limited to rash (3/10) and insomnia (1/10). Microbiological cultures in patients able to cough up sputum (8/10) were comparable to those collected at baseline.

Breath analysis

Detection of volatile organic compounds

A total of 29 breath samples were obtained, from which 99 features were consistently detected. Five groups of correlated features were identified, corresponding to water clusters or dehydrated forms of alcohols or short chain fatty acids (SCFA) [23–25,32]: *m/z* [33.03/51.04] (methanol and water cluster); *m/z* [47.05/65.06] (ethanol and water cluster); *m/z* [61.03/79.04/43.02] (acetic acid; $\pm H_2O$); *m/z* [75.04/57.03] (propionic acid; $-H_2O$); *m/z* [89.06/71.05/107.08] (butyric acid; $\pm H_2O$). Redundant information from these groups was eliminated by keeping only parent ions, which resulted in a total of 92 features available for statistical analysis. Unsupervised PCA demonstrated a homogeneous dataset, with patient identity representing the dominant source of variability (Figure S3 in the Supporting Information).

Changes in breath profiles after ETI initiation

Longitudinal analysis using LMM identified 11 features that changed in responders following ETI initiation (*p*-value <0.05, Table 2, Figure 1). Of these, eight exhibited different levels in pwCF at V0 compared to healthy volunteers (*p*-value <0.05). Modifications in CF breath profiles were detectable from V1 onwards, progressively shifting towards healthy levels. By V2, seven features remained differentially expressed in pwCF.

The RM-ASCA+ model indicated that 5% of the total variance was attributable to time. Scores of the first two principal components (PC) are shown at the top of Figure 1 (with corresponding loadings shown in Figure S4 in the Supporting Information). PC1, which captures 53% of the time-related variance, reflects features displaying linear trajectories over time, either increasing (e.g. *m/z* 85.10) or decreasing (e.g. *m/z* 74.06). In contrast, PC2 (47%) encompasses features with divergent temporal patterns at V1 and V2. It

therefore includes a range of profiles, such as features that shifted at V1 and subsequently returned to baseline (e.g. *m/z* 87.08) or stabilised at V2 (e.g. *m/z* 61.03), as well as features that exhibited minimal change at V1 but substantial modifications at V2 (e.g. *m/z* 63.04).

The multilevel PLS-DA predicted patient visits with an AUC of 0.89 for V0 and V1, and 0.73 for V2 (Figure 2A). Classification performance improved substantially when using a reduced set of 14 selected features, yielding an AUC of 1.00 for V2 and 0.96–0.97 for V0 and V1 (Figure 2B). These 14 features represent the union of the 11 features identified by the LMM and the 10 highest VIPs from the PLS-DA, thereby introducing three additional features (*m/z* 47.05, *m/z* 79.08 and *m/z* 41.04; Table 2, Figure 1).

Correlations with clinical outcomes

A total of 10 breath features were significantly associated with one or more clinical readouts (Table 2, Figure 3, Figure S5 in the Supporting Information). Five features correlated with ppFEV□, seven with ppFVC, and two with SCC. Notably, five of these ten clinically associated features were also significantly modified following treatment in responders.

Feature annotation

In addition to the five previously described clusters, supposedly corresponding to methanol, ethanol, acetic, propionic and butyric acids, several tentative annotations were proposed for the 19 features highlighted by the statistical analyses (Table 2). Features *m/z* 73.06, *m/z* 87.08 and *m/z* 135.11 were annotated as 2-butanone, 2-

pentanone and o-cymene, respectively, based on an in-house library. Although structural isomers cannot be dismissed, *m/z* 69.09 is widely recognised as isoprene. Considering their exact mass, *m/z* 74.06 may correspond to N,N-dimethylformamide, propanamide or hydrated propanenitrile, whilst *m/z* 63.04 may be a water cluster of acetaldehyde [25]. Finally, *m/z* 41.04, *m/z* 55.05, *m/z* 57.07, *m/z* 71.09, *m/z* 85.10, *m/z* 99.12 and *m/z* 139.15 most likely correspond to alkenes and/or fragments of larger compounds [25].

DISCUSSION

This prospective, proof-of-concept study provides the first evidence that ETI rapidly modifies the metabolic composition of exhaled breath in adults with CF and that these short-term alterations tend to evolve towards levels observed in healthy individuals. By combining real-time PTR-MS with a multilevel statistical framework, we identified early and progressive volatilomic trajectories reflecting the biological response to ETI, and we observed significant associations between selected VOCs and clinical readouts. These findings highlight the potential of breath analysis as a non-invasive surrogate biomarker for monitoring response to CFTR modulator therapies.

Disease- and pathogen-specific VOC signatures were previously reported in the breath of pwCF [33–35]. However, the impact of CFTR modulators has remained largely unexplored. Recent cross-sectional investigations have yielded heterogeneous results, partly because the widespread use of ETI blurs the distinction between treatment effects and disease-related metabolic signatures. Whilst Mustafina et *al* did not detect ETI-related VOCs by direct PTR-MS analysis [36], Woollam et *al* identified downregulated long-chain aldehydes using off-site gas chromatography-mass spectrometry (GC-MS) [37]. Previous work had reported some modifications induced by the dual therapy lumacaftor/ivacaftor after three months of treatment [13,38], including reductions in 2-butanone and 2-pentanone [38], which align with our findings on ETI. Our study further demonstrates that such changes emerge within a week of ETI initiation and continue to evolve over the first month. In a longitudinal, pilot study [39], we have previously shown early temporal response to ETI in children using GC-MS, also showing correlations between 2-butanone and lung function tests. Of note, the overlap between our findings and previous GC-MS studies is limited, which most likely reflects

differences in analytical technologies, populations, sampling approaches, and timepoints. Indeed, the metabolite coverage of GC-MS and PTR-MS differs substantially. PTR-MS provides rapid, high-sensitivity quantification of small, highly volatile molecules through soft ionisation, whereas GC-MS better resolves large compounds and structural isomers. These approaches should therefore be viewed as complementary. Importantly, our study is the first to include a healthy control cohort analysed on the same instrument, allowing us to contextualise ETI-induced changes relative to physiological levels.

Two distinct temporal patterns emerged from the RM-ASCA+ model. A first group of VOCs showed linear, sustained modifications over time, suggesting progressive biological evolution and slow adaptation under ETI. A second group of features displayed non-linear trajectories, characterised by either a delayed response at one month, or an acute effect at week one followed by stabilisation or a return to baseline at one month. These divergent patterns may reflect different mechanistic layers of ETI action, possibly involving CFTR-dependent and independent mechanisms, including rapid correction of epithelial ion transport, early inflammatory shifts, changes in mucus rheology or clearance, cell membrane restructuring, and slower microbiota reshaping and metabolic rewiring.

Many of the modified VOCs (including SCFAs, ketones, aldehydes, and small alkenes) can be linked to metabolic pathways relevant to CF pathophysiology. We observed that SCFAs such as acetic or butyric acid were elevated in pwCF at baseline and decreased towards healthy levels after treatment. Their origin may be multifactorial, involving both pulmonary and extra-pulmonary sources. ETI has been shown to modulate the lung [40] and gut [41] microbiota, although previous studies did not report major changes in

faecal SCFAs [41]. In the airways, SCFAs have been detected in sputum, where they seem to influence microbial growth and inflammatory responses [42]. Although their exact immune role is debated [43], their normalisation in breath could therefore reflect both microbiome restructuring and reduced inflammatory burden. Other VOCs, including acetaldehyde, ethanol, isoprene, and selected ketones, point towards energy metabolism. Acetate, acetaldehyde, and ethanol are by-products of the glycolytic pathway, and their non-linear evolution suggests a dual acute and delayed metabolic response to ETI. Isoprene, derived from cholesterol biosynthesis [44], and 2-pentanone were shown to rapidly rise in breath at exercise onset. The decrease of all these metabolites is consistent with a reduction in energy expenditure following ETI initiation. Finally, saturated hydrocarbons, commonly reported in breath [12], are poorly detected by PTR-MS due to weak proton affinity [45]. Several significant features may correspond to short-chain alkenes, likely arising from a mixture of endogenous origins (e.g. lipid decarboxylation [46] or membrane turnover) and analytical effects (in-source fragmentation [47]). Although identification remains putative, these metabolomic signals suggest that ETI rapidly impacts the microbiota and shifts the metabolic balance towards a more energy-efficient. Their consistent modification in responders, unmasked by both longitudinal (LMM) and multivariate (PLS-DA) models, together with correlation with clinical outcomes, underpins the robustness of the highlighted features as potentially relevant indicators for treatment monitoring.

Using multilevel PLS-DA, a set of 14 VOCs accurately classified patient visits, achieving AUC values ≥0.96 at V0 and V1 and reaching perfect discrimination at V2. Moreover, several VOCs correlated with clinical outcomes which builds upon recent findings published by Woollam *et al*. and by our team [37,39]. Ten VOCs correlated

with at least one clinical readout, primarily lung function. Notably, butyric acid was associated with ppFEV□, ppFVC, and sweat chloride, and it also changed significantly in responders, making it a promising candidate surrogate biomarker. These results, obtained from brief breath manoeuvres requiring minimal patient cooperation, indicate early and progressive drug-induced biological changes aligning with clinical improvement and underscore the utility of real-time breathomics for clinical monitoring. This suggests that breath analysis could detect treatment benefit even in cases where spirometry or sweat chloride provide limited discriminatory power; for example, in individuals with preserved lung function, young children, or patients with atypical or delayed responses. As such, breathomics could contribute to more precise evaluation of CFTR modulators, especially in the context of theratyping, compassionate use, or emerging modulators targeting rare mutations.

This study has several strengths, including the use of real-time mass spectrometry, the integration of healthy controls, and the application of complementary multilevel statistical models, which identified consistent patterns. PTR-MS enables point-of-care, non-invasive analysis of breath, without sample preparation. It requires minimal patient cooperation and time, in contrast to other explorations (spirometry, sweat test…). Compared to GC-MS, PTR-MS offers complementary advantages in terms of time efficiency and clinical applicability. PTR-MS eliminates the need for sample collection, transportation, storage and preparation thereby reducing the likelihood of bias, degradation and contamination. It has rapid turn-around time, requires little consumable and does not require qualified personnel.

Some limitations must also be acknowledged. The sample size was modest and heterogeneous. Indeed, few patients in our centre were eligible as most were already on

ETI before study start and our PTR-MS instrument cannot be easily deployed to other centres. The study duration was short, preventing assessment of longer-term stabilisation (as seen with lumacaftor/ivacaftor [13]) or full normalisation of VOCs. Moreover, although PTR-MS offers excellent sensitivity for small volatiles, it lacks chromatographic separation and may miss larger or less volatile metabolites. Controlled multicentre studies incorporating harmonised PTR-MS platforms and parallel GC-MS should strengthen chemical annotation and reproducibility.

In summary, this study demonstrates that ETI induces rapid and measurable changes in breath VOC profiles, that these changes tend to evolve towards healthy levels, and that they correlate with respiratory function. Real-time breath analysis therefore emerges as a promising, non-invasive tool to monitor treatment response in cystic fibrosis. Beyond ETI, this approach could facilitate the evaluation of new CFTR modulators, support theratyping strategies, and potentially extend to non-CF airway diseases involving CFTR dysfunction, such as COPD, asthma, or bronchiectasis. [48,49].

LIST OF ABBREVIATIONS

ANOVA: analysis of variance

AUC: area under the curve

CF: cystic fibrosis

CFTR: cystic fibrosis transmembrane conductance regulator

COVID-19: coronavirus disease

E/N: electric field strength

ETI: elexacaftor, tezacaftor, ivacaftor combination

FEV_1: forced expiratory volume in one second

FDR: false discovery rate

FVC: forced vital capacity

GC-MS: gas chromatography – mass spectrometry

IQR: interquartile range

LOD: limit of detection

m/z: mass/charge ratio

PC: principal component

PLS-DA: partial least-squares – discriminant analysis

pp: percentage of predictive values

PTR-MS: proton-transfer reaction – mass spectrometry

PTR-Qi-TOF MS: proton-transfer reaction – quadrupole – time-of-flight mass spectrometer

pwCF: people with CF

RM-ASCA+: repeated measures ANOVA–simultaneous component analysis plus

rmcorr: repeated measures correlation

ROC: receiver operating characteristic

SCC: sweat chloride concentration

SCFA: short-chain fatty acid

VIP: variable importance in projection

VOC(s): volatile organic compound(s)

DECLARATIONS

Ethics approval and consent to participate

Patients were participating to the PHEAL-KAFTRIO study, which was approved by an ethics committee (Comité de protection des personnes Sud-Est I, 2021-A03119-32, NCT05295524). All patients provided written informed consent.

Data availability statement

All data supporting the findings of this study, including summary breath metabolomics results and clinical outcomes, are included in the main manuscript and Supplementary Information.

Competing interests

ISG reports a Vertex Innovation Award, consulting fees and travel support from Vertex therapeutics. CR and SGD are co-founders of Exhalon. The other authors declare that they have no competing interests.

Funding

This work was supported by Région Île de France (VolatolHom, SESAME 2016 and MeLoMane, DIM 1HEALTH 2019), Fondation Foch (VolatolHom and VOC-info), and financial grants from the European Cystic Fibrosis Society and CF Europe (2020-2022) and Vaincre la Mucoviscidose (RC20220503003).

Authors' contributions

EB, HS, DG, SGD designed and planned the study; HS, SDM, DG, SGD conducted the study; CR, NH led the data analysis; EB, HS, EL, DG, SGD interpreted the results; EB, HS, SGD wrote the manuscript; ISG, PD revised the manuscript; all authors read and approved the final manuscript; SGD is responsible for the content of this study as the guarantor.

Acknowledgments

The authors wish to acknowledge Thao Nguyen Khoa (Cochin hospital, Paris) for SCC analyses; Antoine Bertrand, Pheal company, for assistance in clinical trial promotion; Bedra Bessaidi, Beatrice d'Urso and Flavie Barret for technical investigation work. During the preparation of this work, the authors used an artificial intelligence tool solely to assist with language editing and to improve the clarity of the manuscript. The authors reviewed and validated all content and take full responsibility for the accuracy and integrity of the final submission.

REFERENCES

1 Middleton PG, Mall MA, Dřevínek P, *et al.* Elexacaftor–Tezacaftor–Ivacaftor for Cystic Fibrosis with a Single Phe508del Allele. *N Engl J Med*. 2019;381:1809–19. doi: 10.1056/NEJMoa1908639

2 Allen L, Allen L, Carr SB, *et al.* Future therapies for cystic fibrosis. *Nat Commun*. 2023;14:693. doi: 10.1038/s41467-023-36244-2

3 Aalbers BL, Mohamed Hoesein FAA, Hofland RW, *et al.* Radiological and long-term clinical response to elexacaftor/tezacaftor/ivacaftor in people with cystic fibrosis with advanced lung disease. *Pediatric Pulmonology*. 2023;58:2317–22. doi: 10.1002/ppul.26486

4 Schütz K, Pallenberg ST, Kontsendorn J, *et al.* Spirometric and anthropometric improvements in response to elexacaftor/tezacaftor/ivacaftor depending on age and lung disease severity. *Front Pharmacol*. 2023;14:1171544. doi: 10.3389/fphar.2023.1171544

5 Dreano E, Burgel PR, Hatton A, *et al.* Theratyping Cystic Fibrosis patients to guide Elexacaftor-Tezacaftor-Ivacaftor out of label prescription. *European Respiratory Journal*. Published Online First: 1 January 2023. doi: 10.1183/13993003.00110-2023

6 Fidler MC, Beusmans J, Panorchan P, *et al.* Correlation of sweat chloride and percent predicted FEV1 in cystic fibrosis patients treated with ivacaftor. *Journal of Cystic Fibrosis*. 2017;16:41–4. doi: 10.1016/j.jcf.2016.10.002

7 Heijerman HGM, McKone EF, Downey DG, *et al.* Efficacy and safety of the elexacaftor plus tezacaftor plus ivacaftor combination regimen in people with cystic fibrosis homozygous for the F508del mutation: a double-blind, randomised, phase 3 trial. *The Lancet*. 2019;394:1940–8. doi: 10.1016/S0140-6736(19)32597-8

8 Saluzzo F, Riberi L, Messore B, *et al.* CFTR Modulator Therapies: Potential Impact on Airway Infections in Cystic Fibrosis. *Cells*. 2022;11:1243. doi: 10.3390/cells11071243

9 Lepissier A, Bonnel AS, Wizla N, *et al.* Moving the Dial on Airway Inflammation in Response to Trikafta® in Adolescents with Cystic Fibrosis. *Am J Respir Crit Care Med*. Published Online First: 4 January 2023. doi: 10.1164/rccm.202210-1938LE

10 Bass R, Brownell JN, Stallings VA. The Impact of Highly Effective CFTR Modulators on Growth and Nutrition Status. *Nutrients*. 2021;13:2907. doi: 10.3390/nu13092907

11 Ibrahim W, Carr L, Cordell R, *et al.* Breathomics for the Clinician: The use of volatile organic compounds in respiratory diseases. *Thorax*. 2021;76:514–21. doi: 10.1136/thoraxjnl-2020-215667

12 Drabińska N, Flynn C, Ratcliffe N, *et al.* A literature survey of all volatiles from healthy human breath and bodily fluids: the human volatilome. *J Breath Res.* 2021;15:034001. doi: 10.1088/1752-7163/abf1d0

13 Neerincx AH, Whiteson K, Phan JL, *et al.* Lumacaftor/ivacaftor changes the lung microbiome and metabolome in cystic fibrosis patients. *ERJ Open Res*. 2021;7:00731–2020. doi: 10.1183/23120541.00731-2020

14 Grassin-Delyle S, Roquencourt C, Moine P, *et al.* Metabolomics of exhaled breath in critically ill COVID-19 patients: A pilot study. *EBioMedicine*. 2021;63:103154. doi: 10.1016/j.ebiom.2020.103154

15 Roquencourt C, Salvator H, Bardin E, *et al.* Enhanced real-time mass spectrometry breath analysis for the diagnosis of COVID-19. *ERJ Open Research*. Published Online First: 1 January 2023. doi: 10.1183/23120541.00206-2023

16 Jung YJ, Seo HS, Kim JH, *et al.* Advanced Diagnostic Technology of Volatile Organic Compounds Real Time analysis Analysis From Exhaled Breath of Gastric Cancer Patients Using Proton-Transfer-Reaction Time-of-Flight Mass Spectrometry. *Front Oncol*. 2021;11:560591. doi: 10.3389/fonc.2021.560591

17 Stanojevic S, Vukovojac K, Sykes J, *et al.* Projecting the impact of delayed access to elexacaftor/tezacaftor/ivacaftor for people with Cystic Fibrosis. *Journal of Cystic Fibrosis*. 2020;S1569199320308092. doi: 10.1016/j.jcf.2020.07.017

18 Castellani C, Duff AJA, Bell SC, *et al.* ECFS best practice guidelines: the 2018 revision. *Journal of Cystic Fibrosis*. 2018;17:153–78. doi: 10.1016/j.jcf.2018.02.006

19 Roquencourt C, Grassin-Delyle S, Thévenot EA. ptairMS: real-time processing and analysis of PTR-TOF-MS data for biomarker discovery in exhaled breath. *Bioinformatics*. 2022;38:1930–7. doi: 10.1093/bioinformatics/btac031

20 Roquencourt C, Lamy E, Bardin E, *et al.* A benchmark study of data normalisation methods for PTR-TOF-MS exhaled breath metabolomics. *J Breath Res*. 2024;18:016006. doi: 10.1088/1752-7163/ad08ce

21 Skawinski M, Schooten FJ van, Smolinska A. A comprehensive guide to volatolomics data analysis. *J Breath Res*. 2024;19:015001. doi: 10.1088/1752-7163/ad9b46

22 Kuo T-C, Tan C-E, Wang S-Y, *et al.* Human Breathomics Database. *Database*. 2020;2020:baz139. doi: 10.1093/database/baz139

23 Hartungen E von, Wisthaler A, Mikoviny T, *et al.* Proton-transfer-reaction mass spectrometry (PTR-MS) of carboxylic acids: Determination of Henry's law constants and axillary odour investigations. *International Journal of Mass Spectrometry*. 2004;239:243–8. doi: 10.1016/j.ijms.2004.09.009

24 Sémon E, Arvisenet G, Guichard E, *et al.* Modified proton transfer reaction mass spectrometry (PTR-MS) operating conditions for in vitro and in vivo analysis of wine aroma. *J Mass Spectrom*. 2018;53:65–77. doi: 10.1002/jms.4036

25 Zhang Y, Wang Y, Li C, *et al.* Interpretation of mass spectra by a Vocus proton-transfer-reaction mass spectrometer (PTR-MS) at an urban site: insights from gas chromatographic pre-separation. *Atmospheric Measurement Techniques*. 2025;18:3547–68. doi: 10.5194/amt-18-3547-2025

26 Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*. 2017;82:1–26. doi: 10.18637/jss.v082.i13

27 Jarmund AH, Madssen TS, Giskeødegård GF. ALASCA: An R package for longitudinal and cross-sectional analysis of multivariate data by ASCA-based methods. *Front Mol Biosci*. 2022;9:962431. doi: 10.3389/fmolb.2022.962431

28 Madssen TS, Giskeødegård GF, Smilde AK, *et al.* Repeated measures ASCA+ for analysis of longitudinal intervention studies with multivariate outcome data. *PLOS Computational Biology*. 2021;17:e1009585. doi: 10.1371/journal.pcbi.1009585

29 Westerhuis JA, van Velzen EJJ, Hoefsloot HCJ, *et al.* Multivariate paired data analysis: multilevel PLSDA versus OPLSDA. *Metabolomics*. 2010;6:119–28. doi: 10.1007/s11306-009-0185-z

30 Rohart F, Gautier B, Singh A, *et al*. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Computational Biology*. 2017;13:e1005752. doi: 10.1371/journal.pcbi.1005752

31 Bakdash JZ, Marusich LR. Repeated Measures Correlation. *Front Psychol*. 2017;8. doi: 10.3389/fpsyg.2017.00456

32 Meurs J, Henderson B, van Dun C, *et al.* Exploring the use of exhaled breath profiling for non-invasive monitoring of cognitive functioning in children: a pilot study. *Metabolomics*. 2025;21:75. doi: 10.1007/s11306-025-02277-5

33 Licht J-C, Seidl E, Slingers G, *et al.* Exhaled breath profiles to detect lung infection with Staphylococcus aureus in children with cystic fibrosis. *Journal of Cystic Fibrosis*. 2023;22:888–93. doi: 10.1016/j.jcf.2023.02.010

34 Kos R, Brinkman P, Neerincx AH, *et al.* Targeted exhaled breath analysis for detection of Pseudomonas aeruginosa in cystic fibrosis patients. *Journal of Cystic Fibrosis*. 2022;21:e28–34. doi: 10.1016/j.jcf.2021.04.015

35 Barucha A, Mauch RM, Duckstein F, *et al.* The potential of volatile organic compound analysis for pathogen detection and disease monitoring in patients with cystic fibrosis. *Expert Rev Respir Med*. Published Online First: 19 July 2022. doi: 10.1080/17476348.2022.2104249

36 Mustafina M, Silantyev A, Krasovskiy S, *et al.* Identification of Exhaled Metabolites Correlated with Respiratory Function and Clinical Features in Adult Patients with Cystic Fibrosis by Real-Time Proton Mass Spectrometry. *Biomolecules*. 2024;14:1189. doi: 10.3390/biom14091189

37 Woollam M, Siegel AP, Grocki P, *et al.* Preliminary method for profiling volatile organic compounds in breath that correlate with pulmonary function and other clinical traits of subjects diagnosed with cystic fibrosis: a pilot study. *J Breath Res*. 2022;16. doi: 10.1088/1752-7163/ac522f

38 van Aerde KJ, Ferwerda G, Smolinska A, *et al.* Immune Response and Exhaled Breath Profile Changes after Initiation of CFTR Modulator Therapy in Children with CF. *International Journal of Translational Medicine*. 2024;4:547–58. doi: 10.3390/ijtm4030037

39 Bardin E, Hunzinger N, Lamy E, *et al.* Short-term modification of breathprint by Elexacaftor/Tezacaftor/Ivacaftor in a paediatric cohort. *Journal of Cystic Fibrosis*. Published Online First: 21 January 2025. doi: 10.1016/j.jcf.2025.01.004

40 Robertson JK, Goldberg JB. The impact of cystic fibrosis transmembrane conductance regulator (CFTR) modulators on the pulmonary microbiota. *Microbiology (Reading)*. 2025;171:001553. doi: 10.1099/mic.0.001553

41 Marsh R, Santos CD, Yule A, *et al.* Impact of extended Elexacaftor/Tezacaftor/Ivacaftor therapy on the gut microbiome in cystic fibrosis. *Journal of Cystic Fibrosis*. 2024;23:967–76. doi: 10.1016/j.jcf.2024.05.002

42 Ghorbani P, Santhakumar P, Hu Q, *et al.* Short-chain fatty acids affect cystic fibrosis airway inflammation and bacterial growth. *Eur Respir J*. 2015;46:1033–45. doi: 10.1183/09031936.00143614

43 Tan JK, Macia L, Mackay CR. Dietary fiber and SCFAs in the regulation of mucosal immunity. *J Allergy Clin Immunol*. 2023;151:361–70. doi: 10.1016/j.jaci.2022.11.007

44 King J, Mochalski P, Kupferthaler A, *et al.* Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study. *Physiol Meas*. 2010;31:1169–84. doi: 10.1088/0967-3334/31/9/008

45 Blake RS, Monks PS, Ellis AM. Proton-transfer reaction mass spectrometry. *Chem Rev*. 2009;109:861–96. doi: 10.1021/cr800364q

46 Monteiro RRC, da Silva SSO, Cavalcante CL, *et al.* Biosynthesis of alkanes/alkenes from fatty acids or derivatives (triacylglycerols or fatty aldehydes). *Biotechnology Advances*. 2022;61:108045. doi: 10.1016/j.biotechadv.2022.108045

47 Koss AR, Sekimoto K, Gilman JB, *et al*. Non-methane organic gas emissions from biomass burning: identification, quantification, and emission factors from PTR-ToF during the FIREX 2016 laboratory experiment. *Atmospheric Chemistry and Physics*. 2018;18:3299–319. doi: 10.5194/acp-18-3299-2018

48 Rowe SM, Jones I, Dransfield MT, *et al*. Efficacy and Safety of the CFTR Potentiator Icenticaftor (QBW251) in COPD: Results from a Phase 2 Randomized Trial. *Int J Chron Obstruct Pulmon Dis*. 2020;15:2399–409. doi: 10.2147/COPD.S257474

49 Patel SD, Bono TR, Rowe SM, *et al.* CFTR targeted therapies: recent advances in cystic fibrosis and possibilities in other diseases of the airways. *Eur Respir Rev*. 2020;29:190068. doi: 10.1183/16000617.0068-2019

50 Amann A, Schwarz K, Wimmer G, *et al*. Model Based Determination of Detection Limits for Proton Transfer Reaction Mass Spectrometer. *Measurement Science Review*. 2010;10. doi: 10.2478/v10048-010-0031-5

FIGURE LEGENDS

Figure 1. The top two graphs represent the two first principal components (PCs) of the RM-ASCA+ model, with 95% percentile bootstrap confidence intervals. The lower plots show the longitudinal evolutions in the breath of responder patients after ETI initiation from baseline (V0), to week one (V1) and month one of treatment (V2), compared to levels measured in healthy volunteers, for 14 selected features. These correspond to the union of the 11 LMM significant features (marked *m/z* *; F-test *p*-value <0.05) and the 10 top VIPs of the multilevel PLS-DA model.

Left panel **(A**) shows features with a linear trend, contributing to PC1 of the RM-ASCA+ model; right panel (**B**) shows features with significant difference at V1 (compared to V0 and/or V2) contributing to PC2. **p*-value <0.05; ***p*-value <0.01; ****p*-value <0.001; *****p*-value <0.0001.

Figure 2. Multilevel PLS-DA showing the distribution of breath samples collected from responder patients at baseline (V0), after one week (V1), and one month (V2) of ETI treatment. The top graphs show the scores plots with shaded areas representing 95% confidence intervals, and the bottom graphs show the ROC curves comparing each visit against the others. Left panel (**A**) shows the results of the model using all features (92); right panel (**B**) shows the model built with the 14 selected features.

Figure 3: Repeated measures correlations (rmcorr) between two significant VOCs and ppFEV□ (**A**), ppFVC (**B**), and SCC (**C**). The data for each patient are linked by an arrow, in the order of visits (V0->V1->V2), and the black dotted line represents the estimated correlation. The first row shows the same feature (*m/z* cluster [89.06/71.05/107.08]), whilst the second row shows a

common feature for (**A**) and (**B**) (*m/z* 85.10), and another one (*m/z* 69.07) for (**C**). All significant correlations are shown in Figure S5 in the Supporting Information.

Table 1. Cohort characteristics and clinical readouts during the study.

Patient		**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**	**10**
Sex		F	M	M	M	M	M	M	M	F	M
Age range (y)		26-30	41-45	36-40	31-35	51-55	31-35	21-25	46-50	21-25	66-70
Genotype		F508del/ S1255P	F508del/ F508del	F508del/ F508del	F508del/ S945L	S945L(exon15) / I601F	3120G>A/ 3120G>A	N1303K/ N1303K	F508del/ E92K	F508del/ I148T	F508del/ 3849+10kbC>T
CFTR modulator treatment at V0		Ivacaftor	-	-	-	Ivacaftor	-	-	-	-	-
Therapy access		Eligible	Eligible	Eligible	Eligible	Off-label	Off-label	Off-label	Eligible	Eligible	Eligible
Sweat chloride (mmol/L)	**V0**	15	Failed	92	100	24.5	106	96	92.5	100	64
	V1	<10	ND	17	60	28.5	124	93	38	51.5	32.5
	V2	ND	ND	ND	58.5	<10	100.5	108	21.5	52	44
ppFEV1 (%)	**V0**	35	54	120	80	24	30	18	67	87	76
	V1	38	66	129	91	25	28	20	74	96	78
	V2	43	80	137	94	25	28	21	77	96	87
ppFVC (%)	**V0**	64	61	125	89	63	48	55	88	99	80
	V1	68	94	128	99	51	46	61	91	103	81
	V2	71	103	124	100	52	46	72	94	105	91
Clinical improvement at V2			Cough Dyspnea +3kg	Cough Sputum	Cough Sputum +5kg	Fatigue Sputum +1kg		Dyspnea Sputum +1kg	Cough Dyspnea Sputum	Cough Dyspnea Sputum	Cough Dyspnea +2kg
Response to ETI		Inconclusive	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
ETI interruption		at one month	No	No	No	No	at two months	No	No	No	No

V0: baseline; V1: week 1; V2: month 1 of ETI treatment. $ppFEV_1$: percent predicted in forced expiratory volume in 1 second; ppFVC: percent predicted in forced vital capacity; ND = Not done.

Table 2. VOCs significantly modified upon ETI treatment and/or correlated with variations in clinical readouts.

m/z	**Exhaled breath** (mean, ppb)	**LMM model**			**Correlations**						**Molecular formula**	**Tentative annotations#**
					ppFEV1		**ppFVC**		**Sweat chloride**			
		p-value	β_1	β_2	*p*-value	R	*p*-value	R	*p*-value	R		
41.04	246.94										$[C_3H_4+H]^+$	Fragments
55.05	7.86	0.041	-0.22	-0.06								Fragments
57.07	9.47	0.029	1.16	1.22			0.050	0.49			$[C_4H_8+H]^+$	C4-Alkenes and fragments
63.04	1.07	0.027	0.10	-0.32								Acetaldehyde+H_2O
67.07	0.55	0.016	-0.57	-1.11								
69.09	20.93								0.011	0.62	$[C_5H_8+H]^+$	Isoprene
71.09	2.41						0.003	-0.63				C5-Alkenes and fragments
73.06	1.73						0.032	-0.65				2-Butanone
74.06	0.31	0.038	-0.15	-0.63	0.031	-0.6	0.026	-0.52			$[C_3H_7NO+H]^+$	N,N-dimethylformamide, propanamide, propanenitrile+H_2O
79.08	0.38											
85.10	1.03	0.001	0.47	1.06	0.004	0.68	0.038	0.5			$[C_6H_{12}+H]^+$	C6-Alkenes and fragments
87.08	0.67	0.006	-0.96	-0.23							$[C_5H_{10}O+H]^+$	2-Pentanone
99.12	0.21				0.037	0.49					$[C_7H_{14}+H]^+$	C7-Alkenes and fragments
103.08	0.19	0.053	-0.12	-0.33	0.009	-0.58					$[C_5H_{10}O_2+H]^+$	Pentanoic acid
135.11	0.17						0.043	0.48			$[C_{10}H_{14}+H]^+$	o-Cymene
139.15	0.06	0.027	0.52	-0.88							$[C_{10}H_{18}+H]^+$	C10-Alkenes
89.06/ 71.05/ 107.08	5.38	0.051	-0.27	-0.33	0.024	-0.5	0.004	-0.62	0.035	0.54	$[C_4H_8O_2+H]^+$	Butyric acid cluster
61.03/ 43.02/ 57.03/ 75.04/ 79.04	32.64	0.052	-0.32	-0.17								Acetic acid cluster
47.05/ 65.06	146.45										$[C_2H_6O+H]^+$	Ethanol cluster

Limits of detection (LOD) were calculated based on [50]; *p*-values obtained with a linear mixed model (LMM) built with patients responding to ETI; β_x LMM coefficients indicates the direction of the evolution at visit V1 and V2; coefficients (R) of correlation tests (rmcorr) are provided where significant. #Tentative chemical identification is based on exact mass match with the human breathomics database [22], likelihood of compound ionisation by proton transfer reaction [45], others' published work [23–25] and an in-house VOC library. All *p*-values are adjusted with the FDR correction.

Figure 1

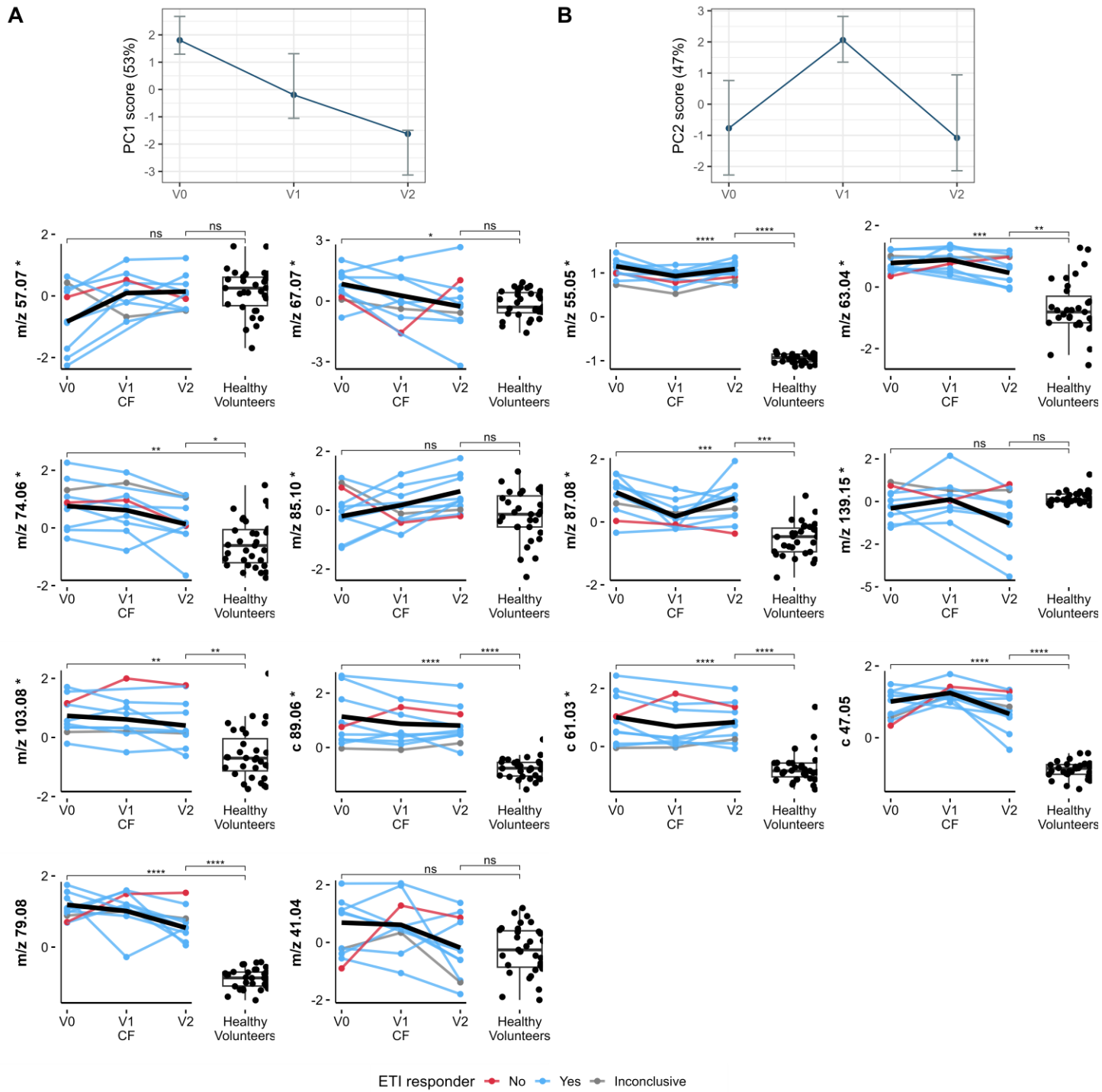

A
PC1 score (53%)
V0
V1
V2
B
PC2 score (47%)
m/z 57.07 *
m/z 67.07 *
m/z 55.05 *
m/z 63.04 *
m/z 74.06 *
m/z 85.10 *
m/z 87.08 *
m/z 139.15 *
m/z 103.08 *
c 89.06 *
c 61.03 *
c 47.05
m/z 79.08
m/z 41.04
V1 CF
Healthy Volunteers
ns
*
**
***
****
ETI responder
No
Yes
Inconclusive

Figure 2

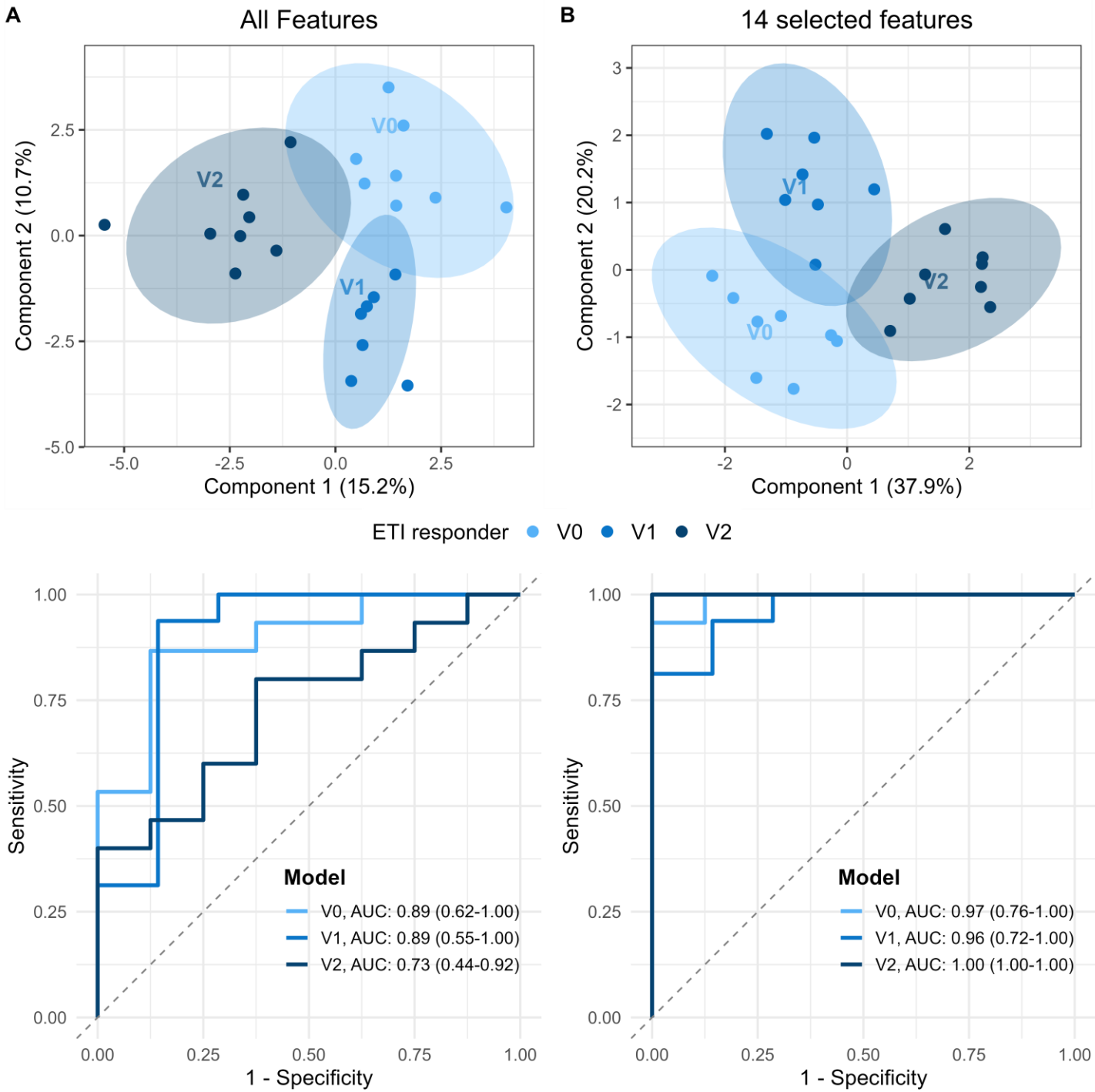
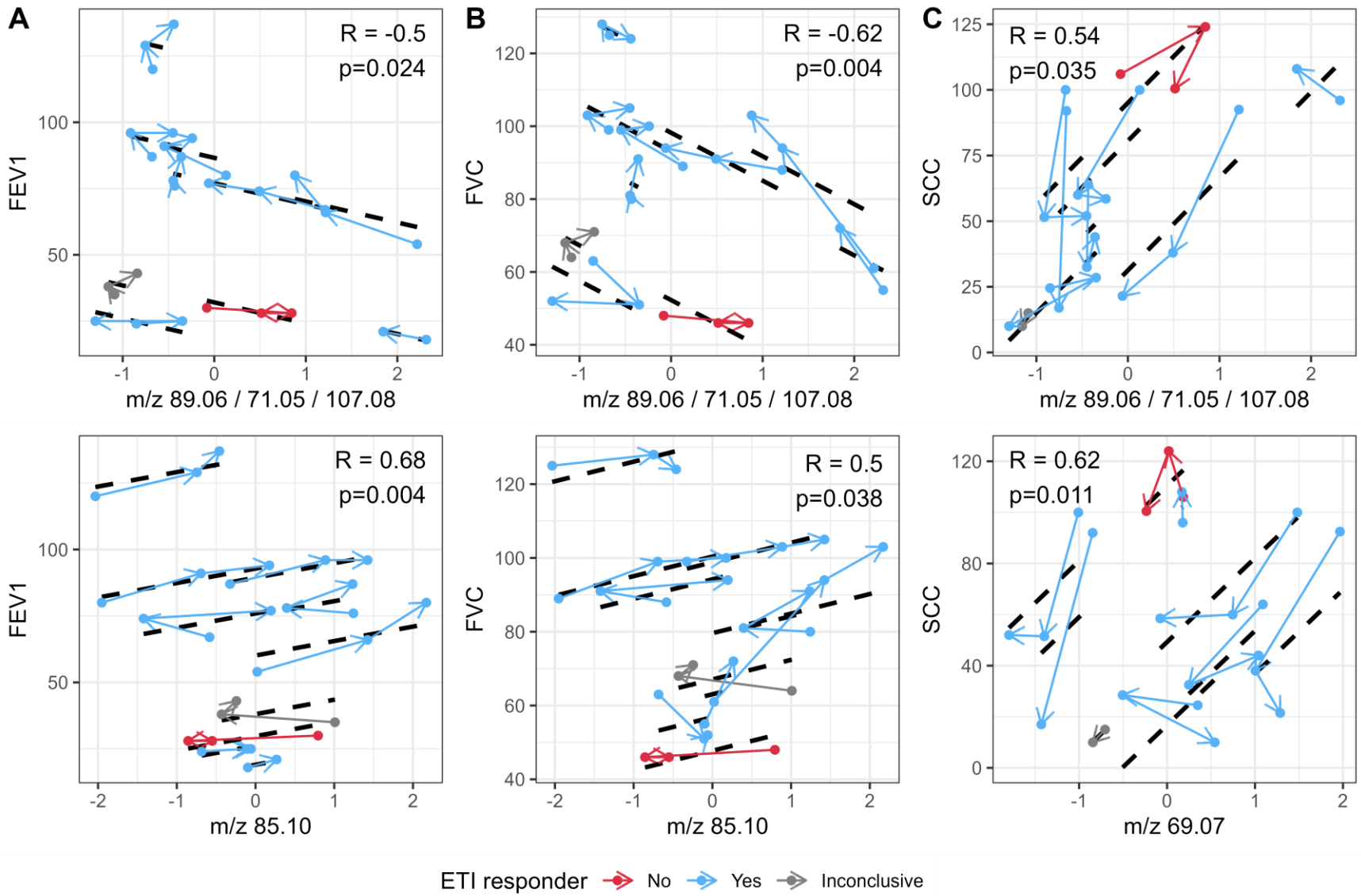

A
All Features
Component 2 (10.7%)
Component 1 (15.2%)
V0
V1
V2
B
14 selected features
Component 2 (20.2%)
Component 1 (37.9%)
V0
V1
V2
ETI responder
V0
V1
V2
Sensitivity
1 - Specificity
Model
V0, AUC: 0.89 (0.62-1.00)
V1, AUC: 0.89 (0.55-1.00)
V2, AUC: 0.73 (0.44-0.92)
Sensitivity
1 - Specificity
Model
V0, AUC: 0.97 (0.76-1.00)
V1, AUC: 0.96 (0.72-1.00)
V2, AUC: 1.00 (1.00-1.00)

Figure 3


A
R = -0.5
p=0.024
FEV1
m/z 89.06 / 71.05 / 107.08
B
R = -0.62
p=0.004
FVC
m/z 89.06 / 71.05 / 107.08
C
R = 0.54
p=0.035
SCC
m/z 89.06 / 71.05 / 107.08
R = 0.68
p=0.004
FEV1
m/z 85.10
R = 0.5
p=0.038
FVC
m/z 85.10
R = 0.62
p=0.011
SCC
m/z 69.07
ETI responder
No
Yes
Inconclusive