

Exhaled Volatile Organic Compounds Are Able to Discriminate between Neutrophilic and Eosinophilic Asthma

Florence N. Schleich¹, Delphine Zanella², Pierre-Hugues Stefanuto², Kirill Bessonov³, Agnieszka Smolinska⁴, Jan W. Dallinga⁴, Monique Henket¹, Virginie Paulus¹, Françoise Guissard¹, Sophie Graff¹, Catherine Moermans¹, Emiel F. M. Wouters⁵, Kristel Van Steen³, Frederik-Jan van Schooten⁴, Jean-François Focant², and Renaud Louis¹

¹Respiratory Medicine, GIGA I³, CHU Sart-Tilman, Liege, Belgium; ²Organic and Biological Analytical Chemistry Group and ³Medical Genomics-BIO3, GIGA-R, University of Liege, Sart-Tilman, Liege, Belgium; ⁴Department of Pharmacology and Toxicology, School for Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, the Netherlands; and ⁵Department of Respiratory Medicine, Maastricht University Medical Center, Maastricht, the Netherlands

Abstract

Rationale: Analysis of exhaled breath for asthma phenotyping using endogenously generated volatile organic compounds (VOCs) offers the possibility of noninvasive diagnosis and therapeutic monitoring. Induced sputum is indeed not widely available and markers of neutrophilic asthma are still lacking.

Objectives: To determine whether analysis of exhaled breath using endogenously generated VOCs can be a surrogate marker for recognition of sputum inflammatory phenotypes.

Methods: We conducted a prospective study on 521 patients with asthma recruited from the University Asthma Clinic of Liege. Patients underwent VOC measurement, fraction of exhaled nitric oxide (F_{ENO}) spirometry, sputum induction, and gave a blood sample. Subjects with asthma were classified in three inflammatory phenotypes according to their sputum granulocytic cell count.

Measurements and Main Results: In the discovery study, seven potential biomarkers were highlighted by gas chromatography–mass spectrometry in a training cohort of 276 patients with asthma. In the

replication study ($n = 245$), we confirmed four VOCs of interest to discriminate among asthma inflammatory phenotypes using comprehensive two-dimensional gas chromatography coupled to high-resolution time-of-flight mass spectrometry. Hexane and 2-hexanone were identified as compounds with the highest classification performance in eosinophilic asthma with accuracy comparable to that of blood eosinophils and F_{ENO}. Moreover, the combination of F_{ENO}, blood eosinophils, and VOCs gave a very good prediction of eosinophilic asthma (area under the receiver operating characteristic curve, 0.9). For neutrophilic asthma, the combination of nonanal, 1-propanol, and hexane had a classification performance similar to F_{ENO} or blood eosinophils in eosinophilic asthma. Those compounds were found in higher levels in neutrophilic asthma.

Conclusions: Our study is the first attempt to characterize VOCs according to sputum granulocytic profile in a large population of patients with asthma and provide surrogate markers for neutrophilic asthma.

Keywords: phenotype; biomarkers; inflammation

(Received in original form November 27, 2018; accepted in final form April 2, 2019)

Supported by the European Respiratory Society, which provided the Fellowship (STRTF 80-2011) to perform the analysis of volatile organic compounds in the discovery study in Maastricht University; the Walloon Region, which provided the Maturation fund, financing the validation study in Liege University; and the Belgian federal government, which provided the EOS 0013618F (ID EOS:30565447) for global financial support.

Author Contributions: Design of the study: F.N.S., E.F.M.W., F.-J.v.S., and R.L. Acquisition of data: F.N.S., M.H., V.P., F.G., S.G., C.M., and R.L. Data analysis and interpretation: F.N.S., D.Z., P.-H.S., K.B., A.S., J.W.D., K.V.S., F.-J.v.S., J.-F.F., and R.L. Drafting the work: F.N.S., D.Z., P.-H.S., E.F.M.W., J.-F.F., and R.L.

Correspondence and requests for reprints should be addressed to Florence N. Schleich, M.D., Ph.D., Respiratory Medicine, GIGA I³, CHU Sart-Tilman B35, 4000 Liege, Belgium. E-mail: fschleich@chuliege.be.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Am J Respir Crit Care Med Vol 200, Iss 4, pp 444–453, Aug 15, 2019

Copyright © 2019 by the American Thoracic Society

Originally Published in Press as DOI: 10.1164/rccm.201811-2210OC on April 11, 2019

Internet address: www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the

Subject: Asthma is a chronic inflammatory airway disease with many different underlying inflammatory mechanisms. Phenotyping asthma according to airway inflammation allows the identification of individuals with asthma who are more likely to respond to targeted therapy. Induced sputum is the gold standard to identify airway inflammatory phenotypes and is currently recommended by guidelines in specialized centers. Several surrogate markers have previously been identified for eosinophilic asthma with acceptable accuracy and there is a lack of biomarkers for neutrophilic asthma.

What This Study Adds to the

Field: We demonstrate that use of four volatile organic compounds (VOCs) was able to discriminate among various types of bronchial inflammation. The combined specificity and sensitivity of VOCs was superior to that of exhaled nitric oxide fraction and blood eosinophil counts that are commonly recognized surrogate markers of sputum eosinophilia. The combination of exhaled nitric oxide fraction, blood eosinophil counts, and VOCs gave a strong prediction model for eosinophilic asthma. VOCs are also able to identify neutrophilic phenotype with accuracy close to that of exhaled nitric oxide fraction, blood eosinophil counts, and VOCs in eosinophilic asthma. This finding is of great interest because markers for neutrophilic asthma are lacking.

Asthma is a chronic inflammatory airway disease with many different underlying mechanisms. Common maintenance treatments of asthma, such as antileukotrienes or inhaled corticosteroids (ICS), and new biotherapies are directed toward modifying the underlying inflammatory process. Phenotyping asthma according to the airway inflammation allows the identification of patients with asthma who are more likely to respond to targeted

therapy. In particular, important studies have confirmed that eosinophilic airway inflammation most reliably predicts the response to antiinflammatory treatment, such as ICS (1, 2) and anti-IL-5. The effect of ICS on innate immunity-driven neutrophilic inflammation is rather poor (3), whereas macrolides (4, 5) seem to confer some benefit in treatment of this subtype (5, 6).

Induced sputum is currently the best available noninvasive technique to assess bronchial inflammation in asthma. Based on this technique, four distinct inflammatory subtypes of asthma have been proposed: 1) eosinophilic, 2) neutrophilic, 3) mixed granulocytic, and 4) paucigranulocytic asthma (7).

Asthma phenotype determination using sputum collection is, however, complex, time-consuming, and not widely applicable. There is a need for new diagnostic methods for characterizing asthma phenotype that are simple, fast, accurate, and cost-effective. Nitric oxide, a type-2 biomarker, may not be the only diagnostically relevant molecule in exhaled breath (8). Breath volatile organic compounds (VOCs) might originate from either inside or outside the body; derived from endogenous products of metabolism or from exogenous sources, such as air, food, and water. Several studies have already suggested the usefulness of VOC detection in exhaled air as a diagnostic tool in some diseases including respiratory disorders, such as tuberculosis, asthma, and chronic obstructive pulmonary disease (9–12). Electronic nose (eNose) has been shown to be able to discriminate between asthma and control subjects with similar accuracy to fraction of exhaled nitric oxide (F_{ENO}) and sputum eosinophils (13). However, eNose identifies a breathprint pattern associated with a disease without precise chemical identity of the compounds present in the exhaled air responsible for the discrimination. Moreover, identification of breath metabolites offers the potential for providing insight into inflammatory pathways (14). Inflammation-associated oxidative stress leads to peroxidation of polyunsaturated fatty acids thereby generating VOCs excreted in exhaled air. We previously identified specific VOCs that are able to discriminate between eosinophilic and neutrophilic inflammation *in vitro* (15).

In this study, we hypothesized that analysis of exhaled breath using endogenously generated VOCs can be a surrogate marker for recognition of sputum inflammatory phenotypes. VOCs offer indeed the possibility of noninvasive diagnosis and therapeutic monitoring.

In the discovery study, seven potential biomarkers for asthma phenotyping were highlighted by gas chromatography (GC)–mass spectrometry (MS) in a discovery training set of subjects with asthma. These VOCs enable the distinction among eosinophilic, neutrophilic, and paucigranulocytic asthma for in-pair comparison scenarios. The following step to assess the diagnostic accuracy using these potential biomarkers relied on the external validation. We tested the diagnostic accuracy of the index test in a new cohort of subjects with asthma, following the Standards for Reporting of Diagnostic Accuracy Studies guidelines (16).

Methods

Subject Characteristics and Study Design

We conducted a prospective cross-sectional discovery study on 276 patients with asthma recruited from the university asthma clinic of Liege between October 2010 and January 2014. In this population, we identified seven compounds as biomarkers discriminating between inflammatory phenotypes. The prospective replication study included 245 patients with asthma who were independently recruited from those in the discovery study, according to the Standards for Reporting of Diagnostic Accuracy Studies guidelines (16). These patients were recruited from the asthma clinic of Liege between October 2014 and April 2018. The recruitment of subjects with asthma stopped when eosinophilic, paucigranulocytic, and neutrophilic phenotypes reached at least 50 subjects with asthma in each group. The mixed granulocytic were excluded from the statistical analysis because of their small occurrence. We used the terms “discovery study” for the whole population recruited to identify VOCs of interest and “replication study” for the whole population used to validate the results of the discovery study. The terms “discovery training set” and “replication training set” were used for the training set of the

discovery and replication studies, respectively. The terms “discovery test set” and “replication test set” were used for the populations of the validation tests of the discovery and replication studies, respectively (Figure 1).

Patients attended the clinic on 2 days at a 1-week interval. On Day 1, each patient underwent VOC measurement, FE_{NO} measurement, spirometry with bronchodilation, sputum induction, gave a blood sample, and filled in validated asthma control and quality-of-life questionnaires (*see the online supplement*). On Day 2, the subjects underwent methacholine challenge after refraining from using bronchodilators for the appropriate time (8 h for short-acting and 24 h for long-acting bronchodilators) as long as the baseline FEV_1 value was not less than 70% predicted.

Asthma was diagnosed according to Global Initiative for Asthma guidelines (17). Patients were classified in four asthma inflammatory phenotypes according to their sputum granulocytic cell count (7). Patients were defined as eosinophilic ($\geq 3\%$ eosinophils in the sputum), neutrophilic ($\geq 76\%$ neutrophils), mixed granulocytic

($\geq 3\%$ eosinophils and $\geq 76\%$ neutrophils), and paucigranulocytic asthma ($< 3\%$ eosinophils and $< 76\%$ neutrophils) (18).

The study was approved by the Ethics Committee of University Hospital of Liege B70720096732, reference Liege 2009/161, and each patient signed an informed consent.

VOC Collection and Analysis

All breath samples were donated between 9:00 A.M. and 11:00 A.M. in the same room, to minimize the effect of variation in background air. Room air was also collected between 9:00 A.M. and 11:00 A.M., and the contribution of ambient room air was constant. Exhaled air was collected by exhaling into inert 5-L Tedlar gas sampling bags. Subjects were asked to inhale, hold their breath for 5 seconds, and subsequently fully exhale into the sampling bag. All sampling bags were washed twice with high-grade nitrogen as described by the manufacturer before usage to make sure all contaminants were eliminated. The content of the sampling bag was transported under standardized conditions onto desorption tubes.

The analysis of the discovery study's samples was conducted in Maastricht

University, where VOCs were analyzed by GC-time of flight (TOF) MS. The replication study samples were analyzed in Liege University using comprehensive two-dimensional GC coupled with high-resolution (HR) TOF MS (GC $\times$ GC-HRTOFMS). GC $\times$ GC offers a significant enhancement of peak capacity over traditional GC. In addition, the measurement of accurate masses enables the identification of target VOCs with confidence. Therefore, GC $\times$ GC-HRTOFMS is an ideal tool for the validation of VOCs in complex samples, such as breath (19). Both analytical procedures have been previously described in detail (20–23). The parameters used for the analysis of samples in the first and in the second prospective studies are listed in the online supplement (*see Table E1 in the online supplement*).

The limit of detection and the linearity range of the seven VOCs were calculated. The concentration range (ppm) of the VOCs of interest is of two orders of magnitude (*see Table E2*). The values of the relative SD at the limit of detection for each standard are reported in Table E3. A quality control procedure, consisting of the

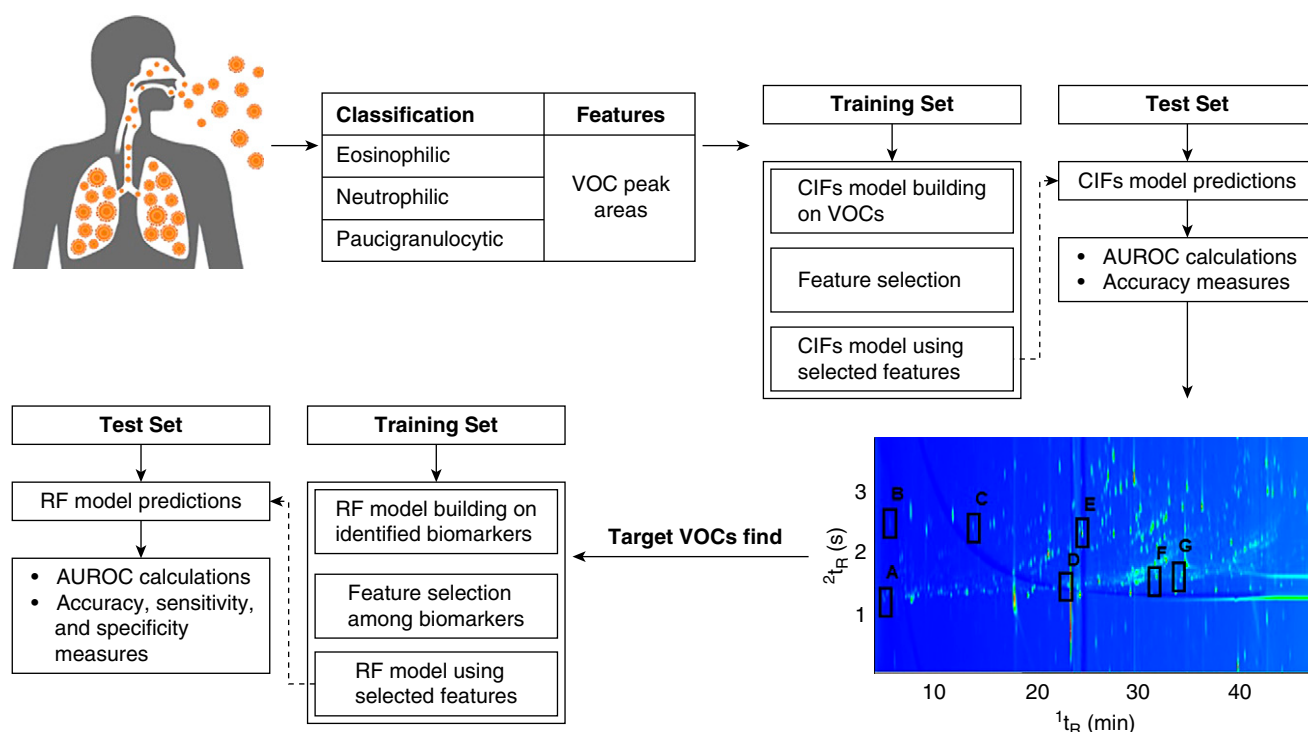


Figure 1. Flow chart of data (pre)processing method and statistical analysis. A = hexane; B = 1-propanol; C = 2-hexanone; D = 3,7-dimethylnonane/undecane; E = nonanal; F = 3-tetradecene; G = pentadecene. ^1t_R = first retention time; ^2t_R = second retention time; AUROC = area under the receiver operating characteristic curve; CIFs = conditional inference forests; RF = random forest; VOC = volatile organic compound.

injection of the standard mixture every 30 samples, was established to monitor the analytical variation (see Figure E1). Analytical standards corresponding to the targeted biomarkers were used (i.e., hexane, 1-propanol, 2-hexanone, nonanal, and undecane) (Sigma Aldrich) to reliably identify the compounds of interest, to establish the linearity range of the analytical instrument, and to monitor the analytical variations of the instrument.

We used a high resolving power analytical approach to efficiently separate the background from the markers. The selection of compounds with high importance value for classification allowed thus to exclude those from the ambient air.

Data Treatment

Analysis of the data output files was performed in successive steps as previously described (21, 24, 25). Data preprocessing, such as baseline and retention time shift corrections, were performed before the creation of the data matrix. In the discovery study, analysis of the data output files was performed in successive steps as previously described (21, 24). Regarding the validation study, standards corresponding to the targeted biomarkers were analyzed. Biomarker identification was performed using chromatographic retention times in the first and second dimensions and their respective mass spectrum. A tolerance of 0.0020 Da, according to the exact mass, was accepted. The biomarkers

were automatically detected in all chromatograms and their raw areas were exported in a spreadsheet for each patient. For both studies, the raw areas of the biomarkers were normalized using probabilistic quotient normalization, before statistical analysis. Probabilistic quotient normalization takes into account the probable dilution factors by analyzing the distribution of the quotients of the signal amplitudes within a test spectrum and comparison with a reference spectrum (20). In this case, the reference spectrum was the calculated median based on all chromatograms of the study.

Statistical Analysis

Statistical analyses were conducted similarly for the biomarker discovery and replication studies. In-pair classifications of phenotypes were investigated: eosinophilic/neutrophilic, eosinophilic/paucigranulocytic, and neutrophilic/paucigranulocytic. To estimate the discrimination importance of each VOC, conditional inference forest algorithm (26) along with variable importance measure were used to assess the discriminatory power of each VOC in the different scenarios. Out of initially selected 12 VOCs from all binary classification sets using the 95th quantile cutoff on the variable importance values (Figure 2), nine VOCs were selected for further clinical validation following minimum performance criteria of accuracy and precision set at 0.60 in individual VOC

classification. Another selection criterion was our ability to map the sequential compound identification to the actual compound, which results in exclusion of two VOCs. Seven VOCs were identified as potential biomarkers for asthma phenotyping. To validate the discrimination importance of the seven biomarkers, random forest algorithm was used. These classification trees are joined together to predict the more accurate feature among a random subset of features (27). For the discovery and the replication studies, the data were split into a training set and a validation set (test set) to assess the classification performances of the biomarkers. Receiver operating characteristic (ROC) curves were constructed to assess the discriminatory power of the biomarkers for asthma phenotyping.

The results of the prospective discovery study were partly previously reported (24).

Results

Patient Characteristics

In the discovery study, 276 subjects with asthma were sampled. From those patients, 122 exhibited eosinophilic asthma, 90 had paucigranulocytic asthma, 50 neutrophilic asthma, and 14 mixed-granulocytic asthma. In the replication study, 245 patients with asthma were recruited, among which were 90 eosinophilic, 53 neutrophilic, 90 paucigranulocytic, and 12 mixed granulocytic. Their demographic functional

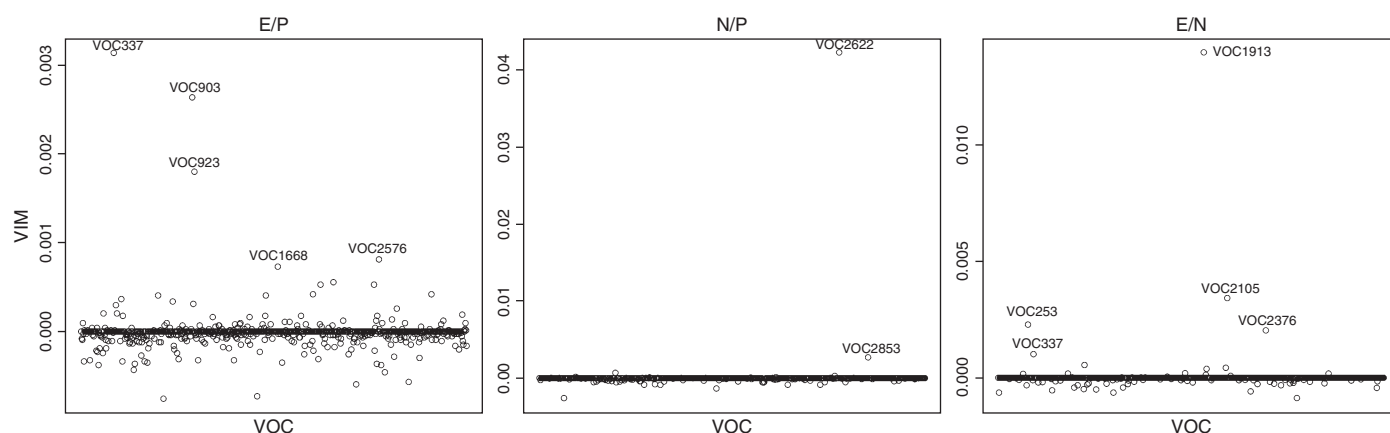


Figure 2. Variable importance measure in eosinophilic versus paucigranulocytic asthma, neutrophilic versus paucigranulocytic asthma, and eosinophilic versus neutrophilic asthma binary classification calculated from conditional inference forests in the discovery study. Totals of five, two, and five volatile organic compounds (VOCs) from the three scenarios were selected as the most important VOCs. Each open circle represents one VOC. Only VOCs providing significant discriminating power are identified with the number reflecting their position on the chromatogram. E/N = eosinophilic versus neutrophilic asthma; E/P = eosinophilic versus paucigranulocytic asthma; N/P = neutrophilic versus paucigranulocytic asthma; VIM = variable importance measure.

Table 1. Demographic and Functional Characteristics of Patients with Asthma Recruited for the VOC Study

Characteristics	Discovery Cohort	Validation Cohort	P Values
N	276	245	
Age, yr	50 ± 15	54 ± 15	0.02
Sex, F, %	59	53	0.19
Height, cm	168 ± 10	168 ± 9	0.35
Weight, kg	73 ± 16	77 ± 16	0.003
Smokers, %	18.5	15	0.20
Exsmokers, %	36	34	0.86
Nonsmokers, %	45.5	50	0.15
Age of onset, %			
<12 yr	22	20	0.26
12–40 yr	36	35	0.30
≥40 yr	42	45	0.03
Atopy, %	60	55	0.50
ICS therapy, yes, %			
Steroid naive	32	41	0.01
Low-dose ICS	15	11	0.18
Moderate-dose ICS	26	17	0.01
High-dose ICS	27	31	0.90
FEV ₁ , % predicted	84 (30–128)	88 (15–148)	0.1
FEV ₁ /FVC, %	70 ± 14	77 ± 14	6 × 10 ^{−4}
Blood eosinophil count, mm ^{−3}	197 (0–1,133)	192 (0–1,547)	0.37
F _{ENO}	26 (6–340)	24 (5–191)	0.08
Sputum eosinophils, %	2.9 (0–91)	2 (0–67)	0.07
Sputum neutrophils, %	50 (3–100)	58 (1–98)	0.0001

Definition of abbreviations: F_{ENO} = fractional exhaled nitric oxide; ICS = inhaled corticosteroids; VOC = volatile organic compound.

Low dose of ICS, <200 µg beclomethasone dipropionate equivalent. Moderate dose of ICS, >200 to 400 beclomethasone dipropionate equivalent. High dose of ICS, >400 µg beclomethasone dipropionate equivalent (Global Initiative for Asthma guidelines) (17). Parametric data are shown as mean ± SD, and nonparametric data are shown as median (range).

and inflammatory characteristics are summarized in Table 1 (for more details, see Tables E4 and E5). The population studied here is a general, unselected population of subjects with asthma including steroid-naïve patients.

Markers of Eosinophilic Asthma

Eosinophilic versus paucigranulocytic asthma. In the discovery study, three VOCs obtained significantly higher variable

importance measure values for the discrimination between subjects with paucigranulocytic and eosinophilic asthma (VOC 337, VOC 903, and VOC 923) (Figure 2, left, and Table 2). The chemical identification of those three VOCs was done by means of spectrum recognition using the National Institute of Standards and Technology (NIST) library in combination with spectrum interpretation by an experienced mass-spectrometrists and identification based on

retention index of compounds. VOC 337 was shown to be hexane, whereas VOC 903 was identified as 2-hexanone (Table 2). VOC 923 remained undetermined and was thus not taken into account in the analysis of the whole population of the replication study. To trace the potential biomarkers highlighted in the discovery phase of this study, it was indeed crucial to get identification of these compounds.

A possible explanation is that this compound is not in the NIST library or that the initial compound structure has been modified during the analysis and is no longer recognizable. We constructed ROC for the eosinophilic versus paucigranulocytic asthma with an area under the curve (AUC) of 0.99 (Figure 3, upper left).

In the replication test set, we confirmed that hexane and 2-hexanone were the compounds with the highest classification performance along with 1-propanol. The AUROC obtained from the ROC curves was 0.68 with an accuracy of 60%, a sensitivity of 60%, and a specificity of 60% (Figure 3, upper right). Hexane and 2-hexanone were found in lower amount in subjects with eosinophilic asthma.

Markers of Neutrophilic Asthma

Neutrophilic versus paucigranulocytic asthma. We further compared the VOCs present in the exhaled air of subjects with neutrophilic as compared with those with paucigranulocytic and eosinophilic asthma to identify discriminative VOCs associated with the neutrophilic phenotype. In the discovery study, we found that VOC 2622 and VOC 2853 were VOCs able to distinguish subjects with asthma with increased neutrophil counts as compared with paucigranulocytic asthma (Figure 2, middle). Using NIST library, we found that VOC 2622 was 3-tetradecene and VOC 2853 was pentadecene (Table 2). 3-Tetradecene was more abundant (8.8 times) and occurred more frequently in neutrophilic asthma. Pentadecene was also present at higher average concentration (4.3 times) and with higher probability of detection in neutrophilic asthma as compared with paucigranulocytic. The ROC curve for the prediction of neutrophilic phenotype using those two VOCs had an AUC of 0.85 (Figure 3, middle left). The caveat is

Table 2. VOC Mappings to Compound Names

VOC ID	Compound Name	Formula
VOC 1913	3,7-Dimethylnonane/undecane	C ₁₁ H ₂₄
VOC 2105	Nonanal	C ₉ H ₁₈ O
VOC2376	Unidentified	
VOC337	Hexane	C ₆ H ₁₄
VOC253	1-Propanol	C ₃ H ₈ O
VOC903	2-Hexanone	C ₆ H ₁₂ O
VOC923	Unidentified	
VOC2622	3-Tetradecene	C ₁₄ H ₂₈
VOC2853	Pentadecene	C ₁₅ H ₃₀

Definition of abbreviation: VOC = volatile organic compound.

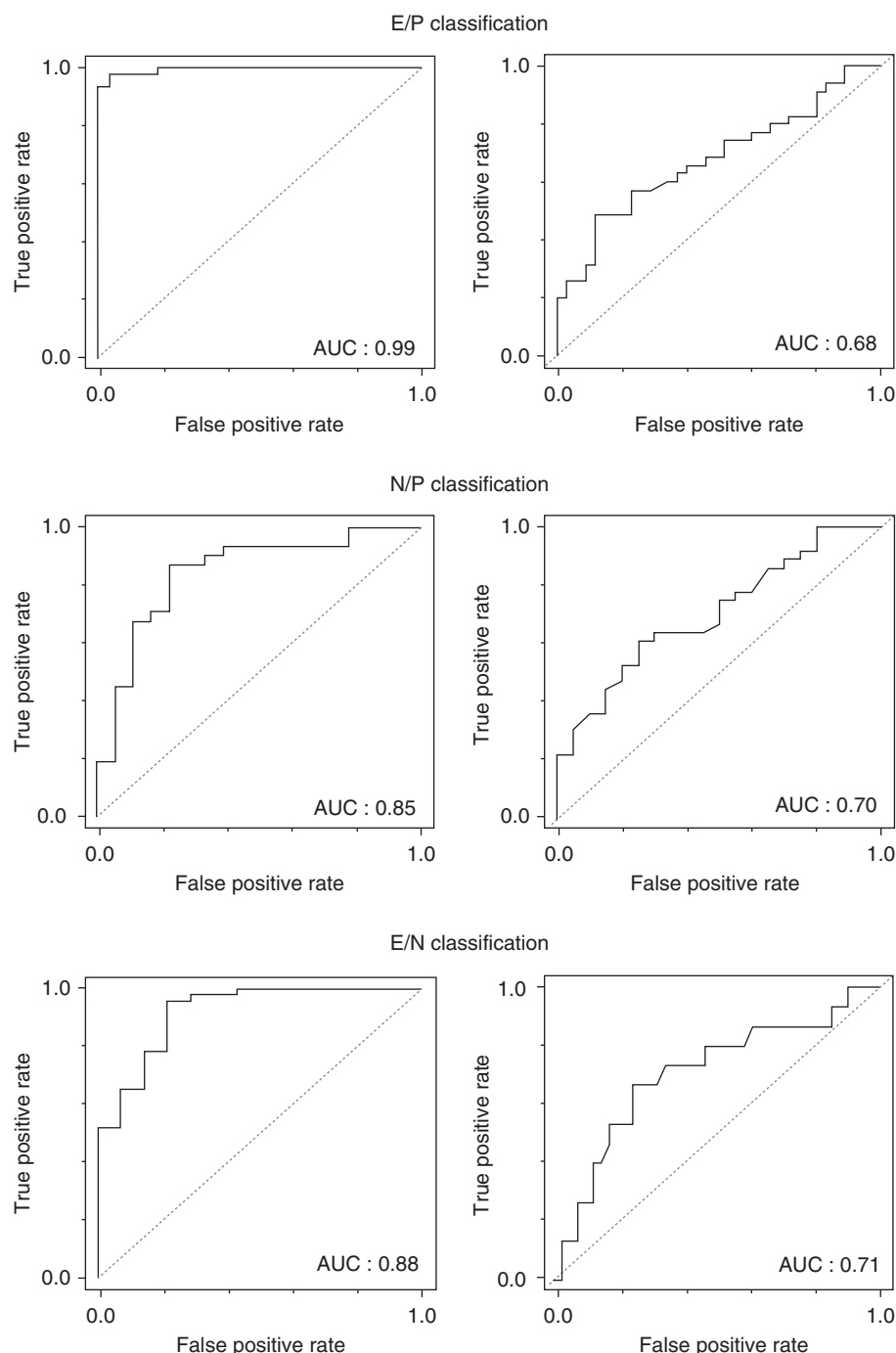


Figure 3. Discovery study on the left side and replication study on the right side. Area under the receiver operating characteristic curve from the conditional inference forest binary classification of eosinophilic versus paucigranulocytic asthma, neutrophilic versus paucigranulocytic asthma, and eosinophilic versus neutrophilic asthma subtypes is shown. The dashed diagonal line indicates random guess, the minimal classifier threshold. Receiver operating characteristic curves from the replication study have a lower area under the curve because biomarkers came from another statistical analysis. AUC = area under the curve; E = eosinophilic asthma subtype; N = neutrophilic asthma subtype; P = paucigranulocytic asthma subtype.

that those two VOCs are not easily transposable from one analytical method to another because of high abundance of isomers in the samples.

In the replication study, the conditional inference forest model highlighted two biomarkers including undecane (isomer of 3,7-dimethylnonane) and nonanal for the

discrimination between paucigranulocytic and neutrophilic asthma (see Figure E2). The classification performance of random forest gave an AUROC of 0.70, an accuracy of 64%, a sensitivity of 71%, and a specificity of 50% (Figure 3, middle right). Nonanal was found in higher amount in subjects with neutrophilic asthma, whereas undecane was found in higher amount in subjects with paucigranulocytic asthma.

Neutrophilic versus eosinophilic asthma. In eosinophilic/neutrophilic classification scenario in the discovery study, we showed that VOC 1913, VOC 2105, and VOC 253 were able to discriminate between eosinophilic and neutrophilic airway inflammation (Figure 2, right). VOC 1913 was identified as 3,7-dimethylnonane in NIST library, whereas VOC 2105 and VOC 253 were found to be nonanal and 1-propanol, respectively. For this cohort the best ranked VOCs were actually 3,7-dimethylnonane and nonanal followed by 1-propanol. The levels of 3,7-dimethylnonane, 1-propanol, and nonanal were on average higher in neutrophilic than in eosinophilic subtype (4.6 times, 3.4 times, and 1.5 times, respectively). We constructed a ROC curve for the classification model between neutrophilic and eosinophilic asthma phenotypes using those three VOCs (AUC, 0.92) (Figure 3, lower left).

In the replication study, by using the combination of these biomarkers in the conditional inference forest model, the combination of hexane and nonanal showed the best classification performances. Therefore, the combination of these two biomarkers was used to build the ROC curve corresponding to the eosinophilic/neutrophilic classification scenario. An AUROC of 0.71 was obtained with an accuracy of 70%, a sensitivity of 45%, and a specificity of 85% (Figure 3, lower right). The amount of hexane, 1-propanol, and nonanal was higher in neutrophilic phenotype, which corroborates the discovery study.

Eosinophilic Asthma versus Other Asthma Phenotypes

The prediction of sputum eosinophilic inflammation in clinical practice relies on the measure of $FeNO$ and blood eosinophil levels. To assess the prediction power of VOCs for sputum eosinophilia, accuracy, sensitivity, and specificity measures were compared with the performance of $FeNO$

Table 3. Accuracy, Sensitivity, and Specificity Measures Observed for Sputum Eosinophilic Inflammation Prediction using F_{ENO} , Blood Eosinophil Count, and VOC Measures

	VOCs	Blood Eosinophil Count	F_{ENO}	Combination
Accuracy, %	72	72	67	76
Sensitivity, %	76	54	44	79
Specificity, %	66	83	83	78

Definition of abbreviations: F_{ENO} = fractional exhaled nitric oxide; VOC = volatile organic compound.

and blood eosinophil levels in the replication study (Table 3). The prediction accuracy using VOC measurement is comparable with the accuracy of blood eosinophil levels and F_{ENO} . However, a better sum of sensitivity and specificity is obtained when using VOCs.

Moreover, using a combination of F_{ENO} , blood eosinophil count, and VOCs gave an AUC of 0.87 (accuracy, 76%; sensitivity, 79%; specificity, 78%) (Figure 4) in the replication study.

Neutrophilic Asthma versus Other Asthma Phenotypes

In the replication test set, the best model to identify neutrophilic asthma included nonanal, propanol, and hexane. The accuracy, sensitivity, and specificity were better combining those three VOCs, and adding undecane, another marker of sputum neutrophilia, did not improve the prediction. The use of nonanal, propanol, and hexane in exhaled air to identify

neutrophilic asthma gave an AUC of 0.73 with 76% accuracy, 81% sensitivity, and 43% specificity (Figure 5).

We selected patients with high sputum neutrophil counts using the following selection criteria: sputum neutrophils greater than or equal to 76% and sputum total cell count greater than or equal to the 95th percentile (i.e., 5.58×10^6 cells/g sputum in a population of 284 healthy subjects recruited in Liege). Among these subjects with highly neutrophilic asthma, using the three VOCs of interest (nonanal, propanol, and hexane), we were able to identify neutrophilic asthma with a higher accuracy (AUC, 0.97; sensitivity, 71%; specificity, 100%) close to the theoretical maximum (Figure 6).

We also looked if age impacted on the VOCs identified because it has been previously shown that the number of sputum neutrophils is increased with ageing. Using the compounds of interest, we did not find any effect of age.

Smoking effect. In a subanalysis of the replication study, it was not possible to discriminate between smokers, exsmokers, and nonsmokers using nonanal, propanol, hexane, and hexanone (accuracy, 50%; sensitivity, 50%; specificity, 46%), suggesting the absence of interference of smoking in our real-life study. Neutrophilic markers were also not associated with smoking pack-years.

ICS effect. We looked at the potential influence of ICS treatment on the association between VOCs and sputum inflammatory profiles. Steroids are indeed changing gene transcription of several inflammatory markers. It was not possible to discriminate between patients treated with ICS and steroid-naïve patients using nonanal, propanol, undecane, hexane, and hexanone (accuracy, 48%; sensitivity, 57%; specificity, 31%; AUC, 0.47), suggesting the absence of interference of ICS therapy in our real-life study.

Discussion

Noninvasive approach of asthma inflammatory phenotypes is of great interest and VOCs were considered as a promising tool (21). In this study, we identified four VOCs of interest to discriminate between asthma inflammatory phenotypes using exhaled air analysis. There is a lack of surrogate markers for neutrophilic asthma. In our study, the combination of nonanal, 1-propanol, and hexane to identify neutrophilic inflammation was similar to the statistical power of an isolated biomarker, such as F_{ENO} or blood eosinophils, to predict eosinophilic asthma. Those compounds were found in higher levels in neutrophilic asthma. For the eosinophilic versus paucigranulocytic asthma discrimination, hexane and 2-hexanone were identified as compounds with the highest classification performance along with 1-propanol, with lower concentration of those VOCs in eosinophilic asthma. The prediction accuracy of sputum eosinophilia using VOCs measurement is comparable with the accuracy of blood eosinophil levels and F_{ENO} . Moreover, the combination of F_{ENO} , blood eosinophil count, and VOCs gave a very good prediction of increased sputum eosinophil counts with an AUROC approaching 0.9.

Based on previous studies, classical GC-MS was used for the discovery phase of the

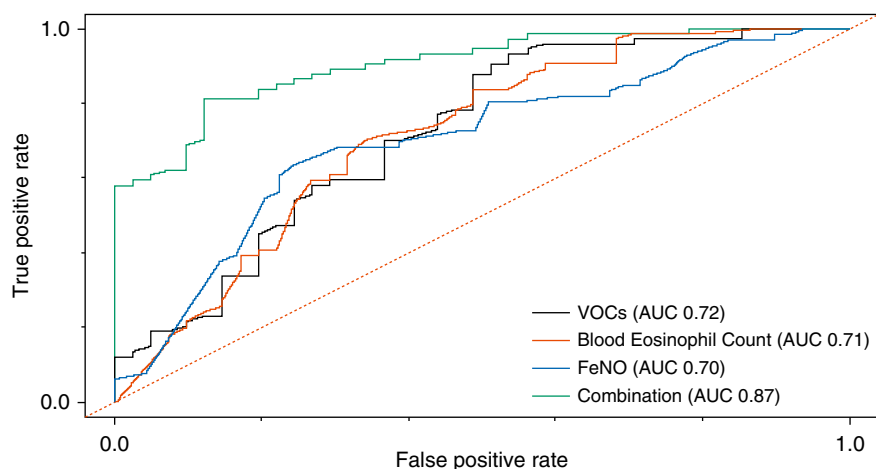


Figure 4. Receiver operating characteristic curve analyses of the sensitivity and specificity of volatile organic compounds (VOCs) (hexane and 2-hexanone), blood eosinophil levels, exhaled nitric oxide fraction (F_{ENO}), and combined VOCs, blood eosinophils, and F_{ENO} for the diagnosis of sputum eosinophilia in the replication cohort. AUC = area under the curve.

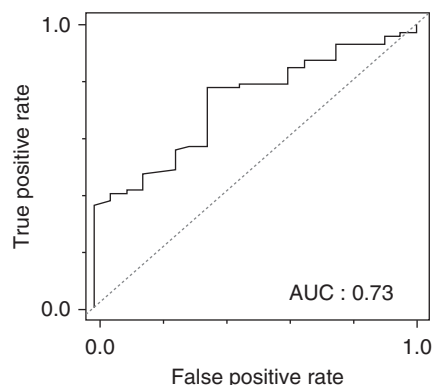


Figure 5. Receiver operating characteristic curve showing the ability of volatile organic compounds (nonanal, 1-propanol, and hexane) to identify neutrophilic asthma in the replication cohort. AUC = area under the curve.

potential biomarkers, whereas GC $\times$ GC-HRTOFMS was used for the replication study because it enables accurate identification of the targeted VOCs along with an increased sensitivity. In previous studies, GC-MS along with eNose have been used to discriminate subjects with asthma versus healthy subjects (22, 25). The ability of eNose to diagnose asthma versus chronic obstructive pulmonary disease was previously demonstrated (23). The study of Montuschi and coworkers (28) has also compared the asthma diagnostic performance of F_{ENO} , lung function testing, and eNose showing that the best performance was obtained by combining eNose and F_{ENO} . Inflammatory

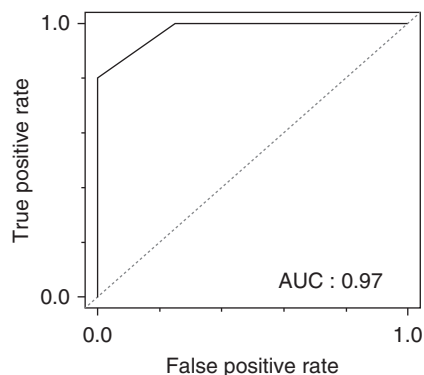


Figure 6. Receiver operating characteristic curve showing the ability of volatile organic compounds (nonanal, 1-propanol, and hexane) to identify neutrophilic asthma in the replication study using highly neutrophilic subjects with asthma (defined by sputum neutrophils $\geq 76\%$ and sputum cell counts greater than the 95th percentile). AUC = area under the curve.

phenotyping studies have been performed using two different systems of eNose (29, 30) using circulating cell numbers as gold standard. Changes in eNose signatures over time were associated with changes in sputum eosinophils in asthma (29). However, eNose does not allow identification of the compounds being surrogate markers of bronchial inflammation. In our study, we confirm the previous observations recently published with eNose that VOCs are able to identify asthma inflammatory phenotypes.

Our study is the first attempt to characterize VOCs according to sputum granulocytic profiles in a large population of subjects with asthma. If we accept the definition of asthma being symptoms combined to excessive airway variability, we believe that it is unlikely that such a complex physiologic disorder may be captured by spotting individual biologic compounds in the exhaled air. By contrast, because granulocytic phenotype may reflect complex inflammatory pathways, we think that VOC analysis may be useful to characterize the inflammatory phenotype of asthma rather than to diagnose asthma itself. Ibrahim and colleagues (31) demonstrated on a small group of 18 subjects with asthma that VOC profiles were able to diagnose sputum eosinophilia using GC-MS. In their study, they created a model using 15 VOCs that was able to discriminate between subjects with asthma and healthy control subjects with 86% accuracy. They also discriminate breath samples from subjects with eosinophilic as compared with noneosinophilic asthma and from subjects with neutrophilic as compared with nonneutrophilic asthma. They found that three compounds had a classification accuracy of 83% to distinguish patients exhibiting induced sputum eosinophil counts less than 2% as compared with greater than or equal to 2% with a sensitivity of 75% and a specificity of 90%. When comparing neutrophilic ($\geq 40\%$) with nonneutrophilic asthma ($< 40\%$), one analyte was able to discriminate between both groups with 72% accuracy. The main limitations of this study were the low number of patients studied and the choice of the neutrophilic threshold, which is well inside the normal range for this cell type (32).

In our large study, we were able to identify neutrophilic phenotype with a better accuracy (76%), whereas VOCs had a

lower accuracy (60%) for eosinophilic phenotype diagnosis. The VOCs highlighted in our study are different from the ones from Ibrahim and colleagues (31) but represent compounds with close chemical structure. 2-Methyldecane ($C_{11}H_{24}$), reported by Ibrahim and coworkers (31), is indeed also an isomer of undecane, found in our study. However, Ibrahim and colleagues (31) did not validate the identification of the compounds using retention indices or using standards corresponding to the targeted biomarkers.

We were able to show on a large population of subjects with asthma that VOCs can be used to discriminate between various types of bronchial inflammation. VOC analysis was able to discriminate between eosinophilic and paucigranulocytic asthma with levels of hexane, 2-hexanone, and 1-propanol being more abundant in paucigranulocytic phenotype. As compared with commonly recognized surrogate markers of sputum eosinophilia, such as F_{ENO} and blood eosinophil counts, the combined specificity and sensitivity of those three VOCs was superior to that of F_{ENO} and blood eosinophil counts. The combination of F_{ENO} , blood eosinophil count, and VOCs gave the best and a very strong prediction model for eosinophilic asthma. The accuracies in predicting sputum cell counts are not perfect and this is not unexpected, because blood granulocyte counts, F_{ENO} , and VOCs represent different biologic compartments than sputum granulocytes.

When comparing VOCs of neutrophilic and the other inflammatory phenotypes of asthma, three VOCs came out as good markers for neutrophilic asthma. Nonanal, 1-propanol, and hexane were indeed found in higher amount in neutrophilic asthmatic exhaled air and were able to identify the neutrophilic phenotype with accuracy close to the one observed to identify sputum eosinophils using F_{ENO} , blood eosinophils, or VOCs. The identification of neutrophilic asthma was close to theoretical maximum using those three VOCs in highly neutrophilic patients defined based on high sputum neutrophils and total sputum cell counts. This finding is of great potential interest because there is a lack of markers for neutrophilic asthma. Indeed, as opposed to what is known for eosinophils (18), blood neutrophils are poor predictor of sputum neutrophils (18, 33).

The VOCs we identified as markers of neutrophilic asthma have been previously reported in other diseases. 1-Propanol was previously shown to be able to discriminate between patients with lung cancer and healthy control subjects (27, 34) and was found in higher levels in patients with lung cancer who exhibit higher numbers of neutrophils in their sputum (35). Nonanal was also found to be elevated in other diseases, such as inflammatory bowel disease (36), gastroesophageal cancer, and breast cancer (37). In the same way, nonanal has been found in higher amounts in exhaled air of smoking healthy volunteers as compared with healthy nonsmokers (38) and cumulative smoking history is well-known to be associated with neutrophilic inflammation (39). Nevertheless, the three VOCs of neutrophilic asthma model were completely insensitive to the smoking history. Neither the categorical classification smokers/exsmokers/nonsmokers nor the cumulative smoking history (pack-years) were indeed associated with increased markers of neutrophils.

In a previous study, steroid usage seemed to be an independent factor in driving eNose-derived asthma phenotypes (29). In our study, it was not possible to discriminate between patients treated with ICS and steroid-naïve patients using VOCs of interest, suggesting the absence of interference of ICS therapy in our real-life study.

In our previous *in vitro* study, we were able to identify VOCs discriminating between eosinophil and neutrophil cultures, whatever the activation status. When taking into account three VOCs or even less to discriminate between the two cellular types, we observed correct classification in the original set and on cross-validation with a high accuracy (15). It is worth noting that none of the compounds we found in the *in vitro* study were identified as discriminating in the *in vivo* study. It might be that VOCs release *in vivo* results from a more complex mechanism than the

spontaneous release observed *in vitro*. In the present study, VOCs might be produced after interaction of eosinophils/neutrophils with resident or structural cells, such as epithelial cells, and could therefore be indirect markers of eosinophilic and neutrophilic inflammation. Moreover, the broad spectrum of biologic responses to and downstream from eosinophilic or neutrophilic inflammation are presumably part of the explanation. It would be interesting to perform an *in vitro* study looking at VOCs released by cocultures of eosinophils/neutrophils and epithelial cells that would maybe be closer to VOCs highlighted in our *in vivo* study.

It is noteworthy to mention that concentration of the VOCs we identified in this study were a thousand times more elevated than FE_{NO} , a well-validated exhaled biomarker in eosinophilic asthma. Whether those VOCs can play a role in asthma pathophysiology remains unknown but we have demonstrated previously that eosinophilic asthma had more severe lung airway obstruction and bronchial hyperresponsiveness (40).

We believe that the four compounds identified may become a powerful biochemical signature to discriminate between inflammatory phenotypes. The construction of the ROC curve for the classification trees containing the most important VOCs were close to the theoretical maximums in the discovery study and those results were confirmed in the replication study using a different analytical method. This supports the robustness of our findings. As expected, the AUC of the ROC curves were lower in the replication as compared with the discovery study. This phenomenon is caused by constraints of the validation. In the discovery study the statistical model to identify compounds that discriminate between the asthma phenotypes has been built taking into account all the compounds identified in the breath of patients with asthma.

Therefore, identifying the compounds of highest importance for in-pair comparison led to AUROC close to the theoretical maximum. While performing the external validation, the compounds that have to be taken into account are imposed to the model. This leads to an expected decrease in the AUC of the ROC curve because the model is built on previously identified biomarkers. In addition, in this study, two different analytical platforms were used to evaluate the classification performances of these biomarkers. However, the model shows that a really good accuracy is obtained for the external replication study.

One potential limitation of our study is the use of sputum cell count, as recommended by current European Respiratory Society/American Thoracic Society guidelines, as the gold standard to identify asthma inflammatory phenotypes. Sputum cell number provides indeed a limited scope on biology and it can be anticipated that complex mixtures of biomarkers, such as VOCs, will eventually better phenotype asthma pathobiology and could become a new gold standard in the future.

Airway inflammatory phenotypes have been shown to be predictive of the response to antiinflammatory treatment. The future for asthma management might rely on the use of small devices to detect VOCs to receive immediate information on inflammatory subtypes. This should help the clinician to adapt immediately their treatment to reach better asthma control with reduction of future risk. Our present work may be a step forward in the process leading to designing exhaled air inflammometers as clinicians use glucometers in diabetes. In a further study, it would be of interest to investigate whether VOCs monitoring might identify responsiveness to maintenance asthma treatments, such as ICSs or macrolides. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1. Green RH, Brightling CE, Woltmann G, Parker D, Wardlaw AJ, Pavord ID. Analysis of induced sputum in adults with asthma: identification of subgroup with isolated sputum neutrophilia and poor response to inhaled corticosteroids. *Thorax* 2002;57:875–879.
2. Pavord ID, Brightling CE, Woltmann G, Wardlaw AJ. Non-eosinophilic corticosteroid unresponsive asthma. *Lancet* 1999;353:2213–2214.
3. Louis R, Schleich F, Barnes PJ. Corticosteroids: still at the frontline in asthma treatment? *Clin Chest Med* 2012;33:531–541.
4. Brusselle GG, Vanderstichele C, Jordens P, Deman R, Slabbynck H, Ringoet V, et al. Azithromycin for prevention of exacerbations in severe asthma (AZISAST): a multicentre randomised double-blind placebo-controlled trial. *Thorax* 2013;68:322–329.
5. Simpson JL, Powell H, Boyle MJ, Scott RJ, Gibson PG. Clarithromycin targets neutrophilic airway inflammation in refractory asthma. *Am J Respir Crit Care Med* 2008;177:148–155.

6. Gibson PG, Yang IA, Upham JW, Reynolds PN, Hodge S, James AL, *et al.* Effect of azithromycin on asthma exacerbations and quality of life in adults with persistent uncontrolled asthma (AMAZES): a randomised, double-blind, placebo-controlled trial. *Lancet* 2017;390: 659–668.
7. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006;11:54–61.
8. Schleich FN, Seidel L, Sele J, Manise M, Quaedyvlieg V, Michils A, *et al.* Exhaled nitric oxide thresholds associated with a sputum eosinophil count $\geq 3\%$ in a cohort of unselected patients with asthma. *Thorax* 2010;65:1039–1044.
9. Robroeks CM, van Berkel JJ, Dallinga JW, Jöbsis Q, Zimmermann LJ, Hendriks HJ, *et al.* Metabolomics of volatile organic compounds in cystic fibrosis patients and controls. *Pediatr Res* 2010;68:75–80.
10. Van Berkel JJ, Dallinga JW, Möller GM, Godschalk RW, Moonen EJ, Wouters EF, *et al.* A profile of volatile organic compounds in breath discriminates COPD patients from controls. *Respir Med* 2010;104: 557–563.
11. Boots AW, Bos LD, van der Schee MP, van Schooten FJ, Sterk PJ. Exhaled molecular fingerprinting in diagnosis and monitoring: validating volatile promises. *Trends Mol Med* 2015;21: 633–644.
12. Kolk AH, van Berkel JJ, Claassens MM, Walters E, Kuijper S, Dallinga JW, *et al.* Breath analysis as a potential diagnostic tool for tuberculosis. *Int J Tuberc Lung Dis* 2012;16:777–782.
13. van der Schee MP, Palmay R, Cowan JO, Taylor DR. Predicting steroid responsiveness in patients with asthma using exhaled breath profiling. *Clin Exp Allergy* 2013;43:1217–1225.
14. Boots AW, van Berkel JJ, Dallinga JW, Smolinska A, Wouters EF, van Schooten FJ. The versatile use of exhaled volatile organic compounds in human health and disease. *J Breath Res* 2012;6:027108.
15. Schleich FN, Dallinga JW, Henket M, Wouters EF, Louis R, Van Schooten FJ. Volatile organic compounds discriminate between eosinophilic and neutrophilic inflammation *in vitro*. *J Breath Res* 2016;10:016006.
16. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, *et al.* STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;6:e012799.
17. Global Initiative for Asthma. Global strategy for asthma management and prevention. 2018 [accessed 2018 Nov 14]. Available from: www.ginasthma.org.
18. Schleich FN, Manise M, Sele J, Henket M, Seidel L, Louis R. Distribution of sputum cellular phenotype in a large asthma cohort: predicting factors for eosinophilic vs neutrophilic inflammation. *BMC Pulm Med* 2013;13:11.
19. Stefanuto PH, Perrault K, Grabherr S, Varlet V, Focant J-F. Postmortem internal gas reservoir monitoring using GCXGC-HRTOF-MS. *Separations* 2018;3:24.
20. Dieterle F, Ross A, Schlöterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures: application in 1H NMR metabolomics. *Anal Chem* 2006;78:4281–4290.
21. Garcia-Marcos L, Edwards J, Kennington E, Aurora P, Baraldi E, Carraro S, *et al.*; EARIP Collaboration. Priorities for future research into asthma diagnostic tools: a PAN-EU consensus exercise from the European Asthma Research Innovation Partnership (EARIP). *Clin Exp Allergy* 2018;48:104–120.
22. Dallinga JW, Robroeks CM, van Berkel JJ, Moonen EJ, Godschalk RW, Jöbsis Q, *et al.* Volatile organic compounds in exhaled breath as a diagnostic tool for asthma in children. *Clin Exp Allergy* 2010;40: 68–76.
23. Fens N, Zwinderman AH, van der Schee MP, de Nijs SB, Dijkers E, Roldaan AC, *et al.* Exhaled breath profiling enables discrimination of chronic obstructive pulmonary disease and asthma. *Am J Respir Crit Care Med* 2009;180:1076–1082.
24. Schleich F. Diagnosis and clinical interest of asthma inflammatory phenotypes [accessed 2018 Nov 14]. Available from: <http://bictel.ulg.ac.be/ETD-db/collection/available/ULgetd-01152015-163140/unrestricted/Schleichthesis.pdf>.
25. Dragonieri S, Brinkman P, Mouw E, Zwinderman AH, Carratù P, Resta O, *et al.* An electronic nose discriminates exhaled breath of patients with untreated pulmonary sarcoidosis from controls. *Respir Med* 2013;107:1073–1078.
26. Hothorn LA. How to deal with multiple treatment or dose groups in randomized clinical trials? *Fundam Clin Pharmacol* 2007;21: 137–154.
27. Ulanowska A, Kowalkowski T, Trawińska E, Buszewski B. The application of statistical methods using VOCs to identify patients with lung cancer. *J Breath Res* 2011;5:046008.
28. Montuschi P, Santonico M, Mondino C, Pennazza G, Mantini G, Martinelli E, *et al.* Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma. *Chest* 2010;137:790–796.
29. Brinkman P, Wagener AH, Hekking PP, Bansal AT, Maitland-van der Zee AH, Wang Y, *et al.*; U-BIOPRED Study Group. Identification and prospective stability of electronic nose (eNose)-derived inflammatory phenotypes in patients with severe asthma. *J Allergy Clin Immunol* 2019;143:1811–1820.
30. de Vries R, Dagelet YWF, Spoor P, Snoey E, Jak PMC, Brinkman P, *et al.* Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. *Eur Respir J* 2018;51:1701817.
31. Ibrahim B, Basanta M, Cadden P, Singh D, Douce D, Woodcock A, *et al.* Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 2011;66:804–809.
32. Louis R, Godinas L, Schleich F. Induced sputum: towards normal values. In: Non invasive assessment of airways inflammation in asthma and COPD. Athens, Greece: Paschalidis Medical Publications; 2011. pp. 113–123.
33. Gibson PG, Simpson JL, Saltos N. Heterogeneity of airway inflammation in persistent asthma: evidence of neutrophilic inflammation and increased sputum interleukin-8. *Chest* 2001;119: 1329–1336.
34. Buszewski B, Ulanowska A, Kowalkowski T, Cieśliński K. Investigation of lung cancer biomarkers by hyphenated separation techniques and chemometrics. *Clin Chem Lab Med* 2011;50: 573–581.
35. Carpagnano GE, Palladino GP, Lacedonia D, Koutelou A, Orlando S, Foschino-Barbaro MP. Neutrophilic airways inflammation in lung cancer: the role of exhaled LTB-4 and IL-8. *BMC Cancer* 2011;11: 226.
36. Hicks LC, Huang J, Kumar S, Powles ST, Orchard TR, Hanna GB, *et al.* Analysis of exhaled breath volatile organic compounds in inflammatory bowel disease: a pilot study. *J Crohn's Colitis* 2015;9: 731–737.
37. Li J, Peng Y, Liu Y, Li W, Jin Y, Tang Z, *et al.* Investigation of potential breath biomarkers for the early diagnosis of breast cancer using gas chromatography-mass spectrometry. *Clin Chim Acta* 2014;436: 59–67.
38. Jareño-Esteban JJ, Muñoz-Lucas MA, Carrillo-Aranda B, Maldonado-Sanz JA, de Granda-Orive I, Aguilar-Ros A, *et al.*; Study Group. Volatile organic compounds in exhaled breath in a healthy population: effect of tobacco smoking. *Arch Bronconeumol* 2013;49: 457–461.
39. Demarche S, Schleich F, Henket M, Paulus V, Van Hees T, Louis R. Detailed analysis of sputum and systemic inflammation in asthma phenotypes: are paucigranulocytic asthmatics really non-inflammatory? *BMC Pulm Med* 2016;16:46.
40. Schleich FN, Chevrement A, Paulus V, Henket M, Manise M, Seidel L, *et al.* Importance of concomitant local and systemic eosinophilia in uncontrolled asthma. *Eur Respir J* 2014;44:97–108.

Exhaled volatile organic compounds are able to discriminate between neutrophilic and eosinophilic asthma

Florence N Schleich¹, Delphine Zanella², Pierre-Hugues Stefanuto², Kirill Bessonov³, Agnieszka Smolinska⁴, Jan W Dallinga⁴, Monique Henket¹, Virginie Paulus¹, Françoise Guissard¹, Sophie Graff¹, Catherine Moermans¹, Emiel FM Wouters⁵, Kristel Van Steen³, Frederick-Jan van Schooten⁴, Jean-François Focant², Renaud Louis¹.

ONLINE DATA SUPPLEMENT

METHODS

Subject characteristics

Asthma was diagnosed based on symptoms of cough, breathlessness or dyspnoea together with the demonstration of airflow variability. The latter was defined by airway hyperresponsiveness demonstrated by one or more of the following: increase in FEV₁ of $\geq 12\%$ and 200 mL following inhalation of 400 μg of salbutamol or inhaled concentration of methacholine provoking a 20% fall in FEV₁ of less than 16 mg/mL. Methacholine challenges were performed according to a standardized methodology as previously described^{E1}. Patients who had a history of upper or lower respiratory tract infection during the 4 weeks before the measurements were excluded from the study. Subjects were characterized as atopic if they had at least one positive specific IgE (> 0.35 KU/L; Phadia) for at least one common aeroallergens (cat, dog, house dust mites, grass pollen, tree pollen, and a mixture of molds). Quality of life was assessed using the self-administered Asthma Quality of Life Questionnaire (AQLQ) and asthma control by the Juniper asthma control questionnaire (ACQ).

FeNO measurement

FeNO was measured by chemoluminescence for the discovery study and by electrochemistry in the validation study, using a nitric oxide monitor set at an exhalation flow rate of 50 ml/sec according to the ERS/ATS recommendations (NIOX Mino – NIOX Vero, Aerocrine, Sweden). FeNO was measured prior to measurement of lung function tests, salbutamol administration, and induced sputum.

Sputum induction and processing

Sputum was induced^{E2} and processed as previously described^{E3}. The differential cell count was performed on cytopins stained with Diff-Quick by counting 500 cells under a light microscope.

VOC collection and analysis

Prior to VOCs collection, all Tedlar bags were washed twice with high-grade nitrogen as described by the manufacturer to make sure all contaminants were eliminated. Table E3 is listing the parameters used for the analysis of samples in the discovery and validation studies.

Table E1. Analytical parameters used for the discovery study (GC-TOFMS) and the validation study (GC×GC-HRTOFMS).

		GC-TOFMS	GC×GC-HRTOFMS
Sampling	Sorption tube	Carbograph 1TD / Carbopack X	Tenax TA / Carbopack B
	Tube desorption	270 °C	290 °C for 5 min
	Injection	Split 10:1	Split 20:1
Analysis			¹ D Rxi-624Sil MS
	Column	RTX-5ms	30 m × 0.25 µm × 1.4 µm df
	combination	30 m × 0.25 mm, 1µm df	² D Stabilwax
			2 m × 0.25 µm × 0.5 µm df
	Oven temperature	40 °C to 270 °C at 10 °C/min	35°C to 240°C at 5 °C/min
Detection	Acquisition range	35 – 350 m/z	29 – 450 m/z
	Acquisition rate	2 Hz	200 Hz
	Ionization energy	70 eV	70 eV

Method performance characteristics of the validation study

For the validation study, the limit of detection, the relative standard deviation (RSD) and the linearity range of the analytical method were calculated according to the EURACHEM guidelines^{E4} by introducing 1 µL of the standards mixture, corresponding to the dilution of the seven VOCs of interest (i.e., hexane, 1-propanol, 2-hexanone, undecane, nonanal, 3-tetradecene and 1-pentadecene) in methanol, in blank stainless steel tubes. Desorption was carried out following the optimized TD-GC×GC-HRTOFMS method.

The concentration range (ppm) of the seven VOCs of interest was evaluated. The response of the instrument obeys a linear relationship for all the targeted compounds whose correlation coefficients are given in Table E4. This working range is compatible

with the concentration of the targeted analytes recovered in breath, making TD-GC×GC-HRTOFMS an appropriate method for biomarkers validation. The values of the RSD at the LOD for each standard are reported in Table E5.

Table E2. Working range of the seven biomarkers targeted in the TD-GC×GC-HRTOFMS validation study.

Compounds	Concentration range (ppm)	Correlation coefficient
Hexane	2 – 10 ppm	0.9876
1-Propanol	2 – 10 ppm	0.9902
2-Hexanone	0.4 – 10 ppm	0.9938
Undecane	0.4 – 10 ppm	0.9866
Nonanal	0.4 – 10 ppm	0.9958
3-Tetradecene	0.4 – 10 ppm	0.9902
1-Pentadecene	0.4 – 10 ppm	0.9867

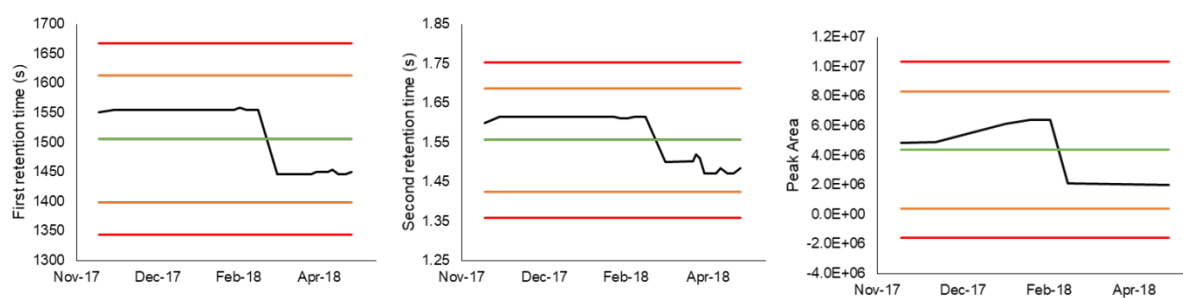
Table E3: Limit of detection (LOD) of the biomarkers and the RSD of the first retention time, the second retention time, and the raw peak area at the LOD.

Compounds	LOD (ppb)	¹ t _R RSD (sec)	² t _R RSD (sec)	Area RSD (%)
Hexane	167	0.8	0.007	2.7
1-Propanol	107	0.0	0.013	3.7
2-Hexanone	2	0.0	0.008	1.8
Undecane	39	2.3	0.009	8.4
Nonanal	30	0.0	0.006	7.9
3-Tetradecene	9	0.0	0.014	4.4
1-Pentadecene	15	0.0	0.003	6.7

Quality control of the validation study

A quality control procedure was established to monitor the analytical changes. The standards mixture was injected every 30 samples. The retention times in both dimension of each analyte was controlled, as well as their raw peak area. Figure E1 is representing the analytical changes of the three different parameters: first retention time, second retention time, and raw area of undecane and 2-hexanone. The green line represents the average, while the orange and red lines represent the 95 % and 99 % confidence level. The shift in both retention times in February corresponds to the column breakage, while the shift in raw area was caused by the maintenance of the ion source. Thanks to these quality control charts, batch effects between samples were monitored and corrected.

A) Undecane



B) 2-Hexanone

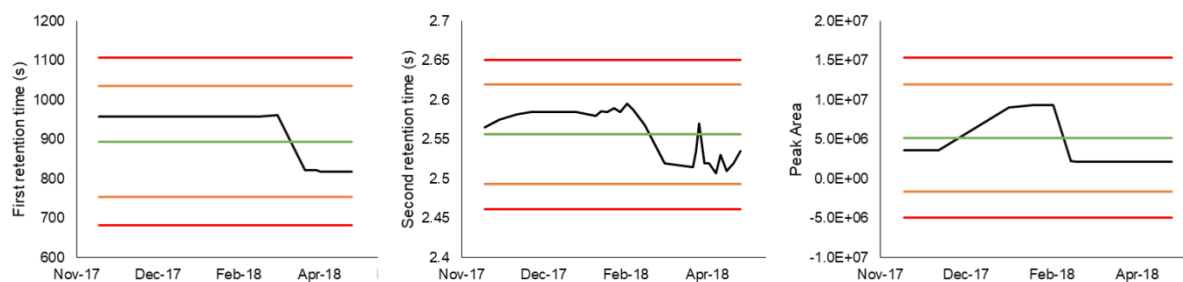


Figure E1. QC charts of undecane and 2-hexanone monitoring, from left to right, the first retention time, the second retention time, and the raw peak area of the compounds.

RESULTS

Patient characteristics

Table E4. Demographic and functional characteristics of the first cohort of asthmatic patients involved in the discovery study according to the inflammatory phenotype.

Characteristics	Eosinophilic Asthma	Neutrophilic Asthma	Paucigranulocytic Asthma
N.	122	50	90
Age (yrs)	53 ± 13***	52 ± 13 ^{°°}	44 ± 17
Gender (% Female)	55	72	58
Smokers (%)	14	25	24
Ex-smokers (%)	39	38	29
Non-smokers (%)	47	37	47
Height, cm	167 ± 13	166 ± 8	169 ± 9
Weight, kg	74 ± 16	72 ± 18	74 ± 18
ICS therapy, Y (%)	67	66	68
Steroid naïve	33	34	32
Low dose ICS	16	14	21
Moderate dose ICS	19	10	16
High dose ICS	32	42	31
FEV ₁ (% pred)	80 (24-133)**	84 (31-124)	91 (30-120)
FEV ₁ /FVC, %	68 ± 14	68 ± 12	75 ± 18
Blood eosinophil count/mm ³	279 (0-1133)***	110 (0-908) ^{μμμ}	127 (0-640)
FeNO	42 (6-248)***	18 (6-280) ^{μμμ}	18 (4-94)
Sputum eosinophils, %	19.2 (3-92)***	0.3 (0-3) ^{μμμ}	0.4 (0-3)
Sputum neutrophils, %	39 (0.7-75)**	86 (76-100) ^{μμμ°°°}	44 (0-76)

Comparison between paucigranulocytic and eosinophilic asthma *p<0.05, **p<0.01, ***p<0.0001. Comparison between neutrophilic and eosinophilic asthma ^μp<0.05, ^{μμ}p<0.01, ^{μμμ}p<0.0001. Comparison between neutrophilic and paucigranulocytic asthma [°]p<0.05, ^{°°}p<0.01, ^{°°°}p<0.0001. Results are expressed in Median (range) or Mean ±SD.

Table E5. Demographic and functional characteristics of the second cohort of asthmatic patients involved in the validation study according to the inflammatory phenotype.

Characteristics	Eosinophilic Asthma	Neutrophilic Asthma	Paucigranulocytic Asthma
N.	90	53	90
Age (yrs)	54 ± 16	59 ± 12	52 ± 16°
Gender (% Female)	50	49	58
Smokers (%)	9**	11	26°°
Ex-smokers (%)	36	42	25°
Non-smokers (%)	55	47	49
Height, cm	168 ± 9	167 ± 11	169 ± 9
Weight, kg	77 ± 15	78 ± 15	77 ± 18
ICS therapy, Y (%)	77	57	41
Steroid naïve	42	39	42
Low dose ICS	11	13	8
Moderate dose ICS	18	13	20
High dose ICS	29	35	30
FEV ₁ (% pred)	87 (19-135)	90 (34-148)	86 (29-119)
FEV ₁ /FVC, %	76 ± 11	79 ± 10	77 ± 11
Blood eosinophil count/mm ³	320 (7 - 1547)***	108 (0 - 691)μμμ	126 (0 - 550)
FeNO	37 (5 - 191)***	17 (4 - 80) μμμ	17 (4 - 90)
Sputum eosinophils, %	14 (3 - 67)***	0.5 (0 - 2.6) μμμ	0.6 (0 - 2.8)
Sputum neutrophils, %	48 (4 - 74)**	88 (77 - 98) μμμ	56 (1 - 76)°°°

Comparison between paucigranulocytic and eosinophilic asthma *p<0.05, **p<0.01, ***p<0.0001. Comparison between neutrophilic and eosinophilic asthma μp<0.05, μμp<0.01, μμμp<0.0001. Comparison between neutrophilic and paucigranulocytic asthma °p<0.05, °°p<0.01, °°°p<0.0001. Results are expressed in Median (range) or Mean ± SD.

Markers of neutrophilic asthma

Permutation test for N/P classification was done using the classification model with other biomarker combinations. The biomarkers' combinations were made randomly. The combination of hexane, 2-hexanone and nonanal made the model collapsing, with an AUROC of 0.47. This shows that only the combination of undecane and nonanal is appropriate for N/P classification. The same results are observed with 1-propanol and nonanal combination, which gave an AUROC of 0.61 (Figure E2).

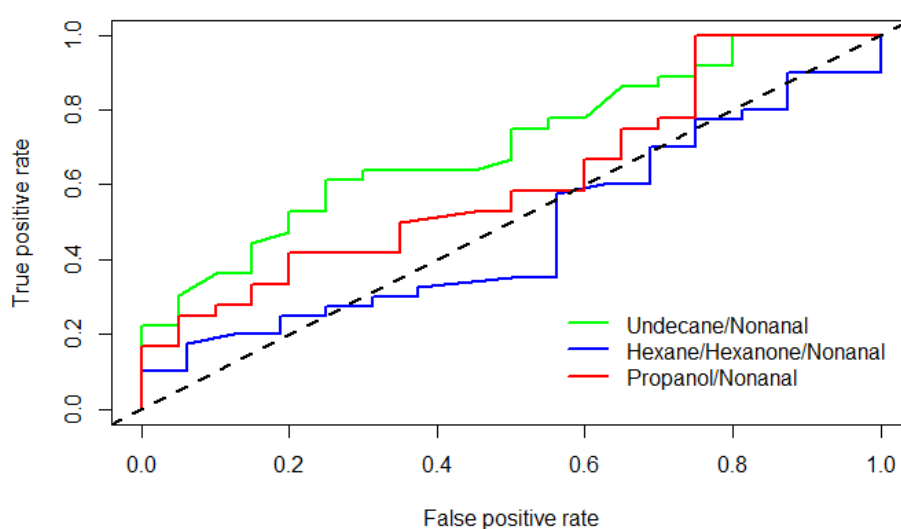


Figure E2. ROC Curves resulting from the permutation test for in-pair classification of the N/P scenario using different combination of biomarkers: undecane/nonanal (Green), hexane/2-hexanone/nonanal (Blue).

Reference List

- E1. Louis, R., J. Sele, M. Henket, D. Cataldo, J. Bettiol, L. Seiden, and P. Bartsch. 2002. Sputum eosinophil count in a large population of patients with mild to moderate steroid-naïve asthma: distribution and relationship with methacholine bronchial hyperresponsiveness. *Allergy* 57:907-912.
- E2. Delvaux, M., M. Henket, L. Lau, P. Kange, P. Bartsch, R. Djukanovic, and R. Louis. 2004. Nebulised salbutamol administered during sputum induction improves bronchoprotection in patients with asthma. *Thorax* 59:111-115.
- E3. Guiot, J., S. Demarche, M. Henket, V. Paulus, S. Graff, F. Schleich, J. L. Corhay, R. Louis, and C. Moermans. 2017. Methodology for Sputum Induction and Laboratory Processing. *J.Vis.Exp.*
- E4. Magnusson B. The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics. 2nd ed. 2014.