



Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label

Rianne de Vries¹, Yennece W.F. Dagelet¹, Pien Spoor², Erik Snoey³, Patrick M.C. Jak⁴, Paul Brinkman¹, Erica Dijkers¹, Simon K. Bootsma⁵, Fred Elskamp⁵, Frans H.C. de Jongh⁶, Eric G. Haarman⁴, Johannes C.C.M. in 't Veen³, Anke-Hilse Maitland-van der Zee¹ and Peter J. Sterk¹

Affiliations: ¹Dept of Respiratory Medicine, Academic Medical Centre, Amsterdam, The Netherlands. ²Faculty of Science and Technology, University of Twente, Enschede, The Netherlands. ³Dept of Pulmonology, Franciscus Gasthuis, Rotterdam, The Netherlands. ⁴Dept of Pediatric Pulmonology, VU University Medical Center, Amsterdam, The Netherlands. ⁵Comon Invent B.V., Delft, The Netherlands. ⁶Dept of Pulmonary Function, Medisch Spectrum Twente, Enschede, The Netherlands.

Correspondence: Rianne de Vries, Dept Respiratory Medicine, F5-260, Academic Medical Centre, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. E-mail: rianne.devries@amc.uva.nl



@ERSpublications

Breathomics may qualify for rapid clinical/inflammatory phenotyping of chronic airway disease at the point of care <http://ow.ly/E16p30gE1Cl>

Cite this article as: de Vries R, Dagelet YWF, Spoor P, *et al.* Clinical and inflammatory phenotyping by breathomics in chronic airway diseases irrespective of the diagnostic label. *Eur Respir J* 2018; 51: 1701817 [<https://doi.org/10.1183/13993003.01817-2017>].

ABSTRACT Asthma and chronic obstructive pulmonary disease (COPD) are complex and overlapping diseases that include inflammatory phenotypes. Novel anti-eosinophilic/anti-neutrophilic strategies demand rapid inflammatory phenotyping, which might be accessible from exhaled breath.

Our objective was to capture clinical/inflammatory phenotypes in patients with chronic airway disease using an electronic nose (eNose) in a training and validation set.

This was a multicentre cross-sectional study in which exhaled breath from asthma and COPD patients (n=435; training n=321 and validation n=114) was analysed using eNose technology. Data analysis involved signal processing and statistics based on principal component analysis followed by unsupervised cluster analysis and supervised linear regression.

Clustering based on eNose resulted in five significant combined asthma and COPD clusters that differed regarding ethnicity (p=0.01), systemic eosinophilia (p=0.02) and neutrophilia (p=0.03), body mass index (p=0.04), exhaled nitric oxide fraction (p<0.01), atopy (p<0.01) and exacerbation rate (p<0.01). Significant regression models were found for the prediction of eosinophilic (R²=0.581) and neutrophilic (R²=0.409) blood counts based on eNose. Similar clusters and regression results were obtained in the validation set.

Phenotyping a combined sample of asthma and COPD patients using eNose provides validated clusters that are not determined by diagnosis, but rather by clinical/inflammatory characteristics. eNose identified systemic neutrophilia and/or eosinophilia in a dose-dependent manner.

This article has supplementary material available from erj.ersjournals.com

Received: June 12 2017 | Accepted after revision: Nov 01 2017

Support statement: This project was sponsored by the public charity Dutch Vriendenloterij. Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: Disclosures can be found alongside this article at erj.ersjournals.com

Copyright ©ERS 2018

Introduction

Asthma and chronic obstructive pulmonary disease (COPD) are now recognised to be complex and heterogeneous chronic airway diseases that include several phenotypes with different inflammatory profiles and responses to treatment. Despite global efforts that are successfully raising the standards of clinical care for asthma [1] and COPD [2], the heterogeneity of these diseases continues to cause difficulties in differentiating the disorders from each other, which leads to uncertainties in management and prognosis of individual patients.

Framing patients based on the historical diagnostic labels of asthma and COPD may not suffice for understanding an individual's condition and for tailoring therapy. Therefore, it has been postulated to gradually move away from categorising patients using such classical disease terminology and to initiate a more personalised approach to management that identifies "treatable traits" [3]. This is underscored by the emergence of highly specific therapies, pointing towards the selection of potential responders based on clinical/inflammatory phenotyping [3, 4]. Management strategies that aim to minimise eosinophilic airway inflammation in both asthma and COPD using inhaled steroids have shown to be effective in decreasing exacerbations [5]. Furthermore, selection of patients based on eosinophilia or neutrophilia appears to increase the efficacy of novel biologicals, such as anti-interleukin-5, or macrolide therapy [3, 5–10].

Inflammatory phenotyping in patients with chronic airway disease can be accomplished by blood or sputum analysis [6]. These assessments do not provide results during patient consultations because they require laboratory procedures. Hence efforts are ongoing to bring biological assessment to the point of care. If properly validated, molecular profiling of exhaled air may provide a noninvasive and rapid alternative for blood and sputum [11]. Volatile organic compounds (VOCs) in exhaled air, obtained by gas chromatography mass spectrometry (GC-MS) have been associated with inflammatory profiles in asthma [12] and COPD [13]. However, identification of individual molecular compounds using GC-MS is not possible or suitable yet for phenotyping in the doctor's office. In contrast, eNose technology can be applied for pattern recognition of the complete mixture of VOCs in exhaled breath without the identification of the individual components [14, 15]. The eNose uses multiple cross-reactive sensors, which means that almost all the individual sensors respond to each analyte, but with different sensitivities. The combined signal response from all the sensors produce a characteristic profile of the VOCs present in exhaled air. This technology is fast and easy to use and has shown great potential to proceed to become a point-of-care tool [16, 17].

We hypothesised that clinical/inflammatory profiles can be obtained from exhaled air in chronic airway diseases, regardless of the diagnosis of asthma or COPD, and that this can be accomplished using eNose technology. Data analysis comes in two main forms: unsupervised (also known as descriptive) and supervised (also known as predictive). Both categories encompass functions capable of finding different hidden pattern in datasets. Therefore, this study purposely used a comprehensive analysis plan. First, unsupervised cluster analysis was performed in patients with chronic airway disease in order to identify

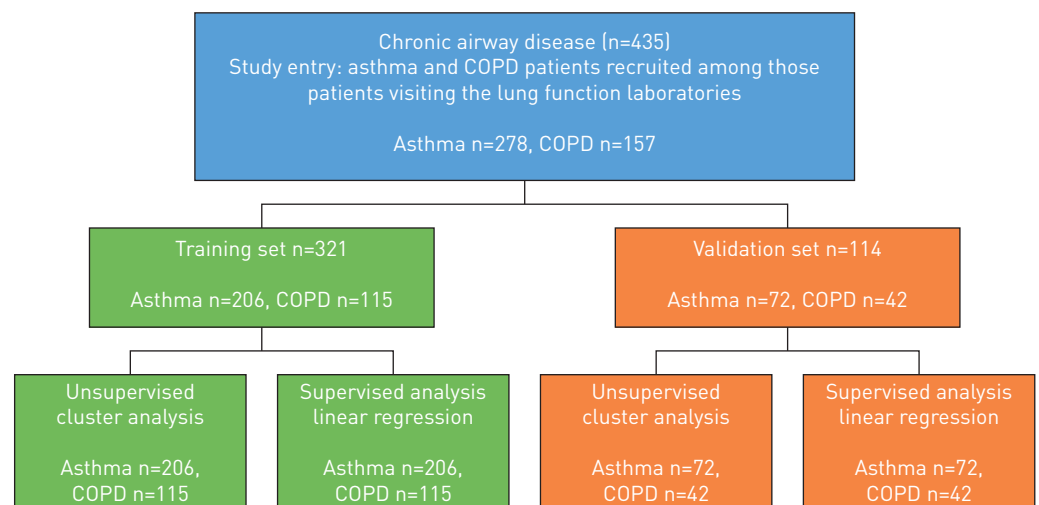


FIGURE 1 Flowchart for patient inclusion in the training and validation set and the design of the comprehensive analysis plan. Unsupervised cluster analysis was used to identify phenotypes of patients based on eNose and supervised analysis was performed to determine the relation between exhaled volatile organic compounds and different inflammatory markers. COPD: chronic obstructive pulmonary disease.

phenotypes of patients based on breathomics. Second, supervised regression analysis was applied to determine the relationship between exhaled VOCs and different inflammatory markers. Our final aim was to test the hypothesis in a validation set of newly recruited asthma and COPD patients.

Methods

Study population

BreathCloud is an ongoing multicentre cross-sectional study (www.breathcloud.org). The data reported in this article include results from the first 435 patients with asthma or COPD, of which 321 patients were used for the training set (asthma $n=206$, COPD $n=115$) and 114 patients for the validation set (asthma $n=72$, COPD $n=42$) (figure 1). Asthma patients were defined by the Global Initiative for Asthma criteria [18] and COPD patients were characterised according to Global Initiative for Chronic Obstructive Lung Disease criteria [2]. The ethics board of all participating sites concluded in writing that Dutch legislation on human participation in research was not considered to be applicable, given the non-invasive nature of this study which merely added exhaled breath analysis to standard diagnostic procedures. Other clinical data (e.g. lung function and eosinophil and neutrophil blood counts) used in this study were collected for routine clinical practice. Given the nature of the study the institutional ethics review board provided a waiver for ethics approval (online supplementary material).

Measurements

Clinical assessment of asthma patients was made using the Asthma Control Questionnaire (ACQ) [19] and for COPD patients the Clinical COPD Questionnaire (CCQ) [20] was used. The personal best post-bronchodilation forced expiratory volume in 1 s (FEV₁; % predicted) was used from data collected for routine clinical practice <12 months prior to the study. Atopy was defined as IgE >0.35 kU·L⁻¹ for at least one of the specific IgEs. Exhaled nitric oxide fraction (FeNO) and eosinophil and neutrophil blood counts were used from clinical data obtained <3 months prior to inclusion. Finally, the presence of exacerbations in the past 3 months was explored. An exacerbation was defined as a worsening of respiratory symptoms which required treatment with oral corticosteroids, antibiotics or both.

Exhaled breath analysis

Exhaled breath analysis was performed in duplicate using the SpiroNose, a validated integration between routine spirometry and eNose technology connected in series with each other [17] (figure 2). Patients were asked to rinse their mouth thoroughly with water three times. Subsequently, exhaled breath analysis was performed in duplicate with a 2-min interval. All patients were instructed to perform five tidal breaths followed by a single inspiratory capacity manoeuvre up to total lung capacity, a 5-s breath-hold and slow (<0.4 L·s⁻¹) maximal expiration towards residual volume (online supplementary figure 1E). Exhaled breath was measured in real time by the SpiroNose sensors. The obtained sensor data were sent directly to and stored at the online BreathCloud server. A more detailed description of the SpiroNose measurement setup and verification of the sensor stability is provided in the online supplementary material.

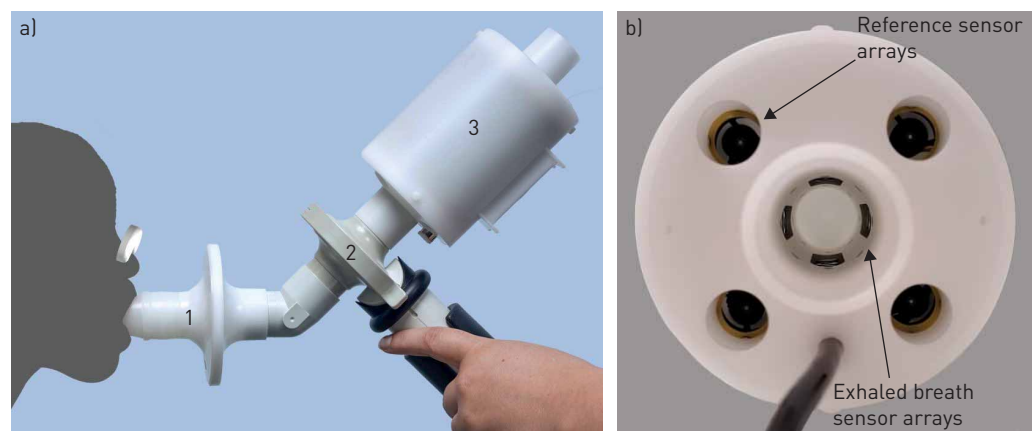


FIGURE 2 a) SpiroNose measurement setup: 1) mouthpiece, nose clamp and bacteria filter; 2) spirometer; 3) SpiroNose. b) Front view of the SpiroNose and the positioning of the eight sensor arrays. The four central sensor arrays monitor exhaled breath; the four reference sensor arrays monitor ambient volatile organic compounds.

Data processing

The offline analysis of raw sensor data downloaded from the BreathCloud server was performed using Matlab (MathWorks, Natick, MA, USA). Details on the processing of the raw sensor signals and the selection of parameters (sensor peak and peak/breath-hold ratios) are explained in online supplementary figure 1E.

Sample size calculation

A sample size calculation was performed based on pilot data from our group. This pilot aimed to determine whether the eosinophil and neutrophil blood counts, measured in a combined sample of asthma and COPD patients, were related to the eNose data. Previous data showed a standard deviation of $\sigma_x=0.42$ and $\sigma_x=0.36$ for eosinophils and neutrophils, respectively. The eNose data (principal component (PC)1) had a standard deviation of $\sigma_y=0.71$. We performed a sample size calculation assuming a slope of $\lambda_1=0.15$ with $\alpha=0.05$ and $\beta=0.2$; the predicted sample size was 180 for the training sample. A validation set half the size of the training set was included, according to recommendations on design and analysis of metabolomics studies [21].

Statistical analysis

A principal component analysis (PCA) was performed to merge the variables of interest into a multivariate component. According to the Kaiser criterion, all PCs with an eigenvalue >1 were retained [17, 22, 23]. The processed sensor variables; the original 14 sensor peaks and the 14 peak/breath-hold ratios (28 variables in total) were restructured to four PCs that captured 91.2% of the variance within the dataset (PC1 46.2%, PC2 21.3%, PC3 13.1% and PC4 10.6%). PCs were constructed for all subjects (training and validation set), based on eNose data from subjects within the training set. Reducing the dimensionality of the data prior to clustering algorithms reduces the risk of overfitting.

Unsupervised analysis

Unsupervised Hierarchical clustering was performed in the combined sample of asthma and COPD patients, on standardised data using Euclidean distance and Ward linkage. We used similarity profile analysis for determining the number of significant clusters [24]. To ensure repeatability and stability of the model, the algorithm was repeated 10 times within random subpopulations (70%) of the training set. Between-cluster comparisons were performed using ANOVA, Kruskal-Wallis or Chi-squared tests. Statistical validity of the results were examined by the identification of eNose clusters within the independent validation set. Finally, PCA plots were generated to present the obtained breath clusters graphically.

Supervised analysis

A multiple linear regression, with stepwise selection of the independent variables was calculated to predict eosinophil blood counts and neutrophilic proportions, measured in the combined sample of asthma and

TABLE 1 Baseline characteristics of the training set

	Asthma	COPD
Subjects	206 [64]	115 [36]
Age years	54.3 \pm 17.1*	65.2 \pm 13.4*
Female/male %/%	59.6/40.4	52.3/47.7
BMI kg·m⁻²	27.1 \pm 5.2	28.2 \pm 6.3
Post-bronchodilator FEV₁ % predicted	83.2 \pm 17.6*	65.9 \pm 21.6*
Smoking (never/ex/current)	124/60/22	1/97/17
Smoking pack-years	8.8 \pm 8.2*	28.4 \pm 11.7*
GOLD (II/III/IV)	NA	56/37/22
GINA (mild/moderate/severe)	63/108/35	NA
Eosinophil blood counts $\times 10^9$ cells·L⁻¹	0.39 \pm 0.41*	0.29 \pm 0.28*
Neutrophil blood counts $\times 10^9$ cells·L⁻¹	5.41 \pm 2.98*	7.15 \pm 3.02*
ICS	175 [85]	81 [70]
OCS	23 [11]	12 [10]

Data are presented as n (%), mean \pm SD or n, unless otherwise stated. COPD: chronic obstructive pulmonary disease; BMI: body mass index; FEV₁: forced expiratory volume in 1 s; GOLD: Global Initiative for Chronic Obstructive Lung Disease; GINA: Global Initiative for Asthma; ICS: inhaled corticosteroid; OCS: oral corticosteroid; NA: not applicable. *: p<0.05.

COPD patients, based on the eNose data (PC1–4). The dependent variables were eosinophil blood counts and neutrophilic proportions measured in blood; the independent variables were the four PCs. A p -value <0.05 was considered statistically significant. Finally, the regression model obtained from the training set was examined in the independent validation set and compared based on the amount of variance that the model accounted for (R^2) and the obtained regression slopes.

Clinical and biological determinants on the eNose data

A multiple linear regression was calculated in the training set to predict eNose data from the key clinical metadata. More information about the analysis and the results is presented in online supplementary tables 3E and 4E.

Results

321 patients were included in the training set. Baseline characteristics showed a significant difference in age, lung function, smoking and systemic eosinophilia and neutrophilia between the asthma and COPD patients (table 1). Asthma patients exhibited significantly higher eosinophil counts ($0.39 \pm 0.41 \times 10^9$ cells·L⁻¹) compared to the COPD group ($0.29 \pm 0.28 \times 10^9$ cells·L⁻¹), whereas neutrophil blood counts were higher in COPD patients ($7.15 \pm 3.02 \times 10^9$ cells·L⁻¹) compared to the asthma patients ($5.41 \pm 2.98 \times 10^9$ cells·L⁻¹).

The validation set consisted of 114 patients. The baseline characteristics (online supplementary table 1E) showed a significant difference in age, lung function, smoking, eosinophilic blood counts and inhaled corticosteroid (ICS) use between the asthma and COPD groups.

Unsupervised analysis

Unsupervised Ward clustering in the training set (T clusters) resulted in five significantly different eNose-driven clusters that differed regarding ethnicity ($p=0.01$), systemic eosinophilia ($p=0.02$) and systemic neutrophilia ($p=0.03$), body mass index (BMI) ($p=0.04$), $FeNO$ ($p<0.01$), atopy ($p<0.01$) and exacerbation rate per person in the past 3 months ($p<0.01$) (table 2 and figure 3). The cluster structure was reproducible when repeating the algorithm within randomly selected subsets of the population. Cluster 1T ($n=74$) contained both asthma (64.9%) and COPD (35.1%) patients who were predominantly females (83.5%) with a high BMI (29.6 ± 5.2 kg·m⁻²), high symptom scores (ACQ 2.87 ± 5.5 , CCQ 4.98 ± 2.2), low $FeNO$ (23.1 ± 7.8) and no inflammation measured in blood (eosinophil counts $0.23 \pm 1.1 \times 10^9$ cells·L⁻¹, neutrophil counts $4.98 \pm 2.2 \times 10^9$ cells·L⁻¹). Cluster 2T ($n=48$) was a combined cluster of asthma (65.1%) and COPD (34.9%) patients who were predominantly males (61.2%) with high circulating eosinophil counts ($0.69 \pm 1.2 \times 10^9$ cells·L⁻¹), high $FeNO$ (48.6 ± 11.0) and low use of oral corticosteroids (OCS) (4.2%). Cluster 3T ($n=24$) contained both asthma (60%) and COPD (40%) patients who were predominately non-Caucasian (90%) with poor lung function (postbronchodilator FEV₁ 68.9% pred) and eosinophil blood counts of $0.45 \pm 1.3 \times 10^9$ cells·L⁻¹. Patients had the lowest exacerbation rate in the past 3 months (0.04 ± 0.19) and no OCS use; the use of ICS (55%) was lowest within this cluster. Cluster 4T ($n=118$) was a combined cluster of asthma (61%) and COPD (39%) patients who were predominantly atopic (84.9%), with high circulating neutrophil ($8.44 \pm 3.1 \times 10^9$ cells·L⁻¹) blood counts. Patients in this cluster experienced the highest number of exacerbations per person in the past 3 months (0.43 ± 0.3). Cluster 5T ($n=57$) had fewer COPD (26.3%) patients, the best postbronchodilator FEV₁ ($83.9 \pm 22.8\%$ pred), an exacerbation rate per person in the past 3 months of 0.14 ± 0.34 and symptom scores that were relatively low (ACQ 1.73 ± 0.9 and CCQ 2.73 ± 1.3).

Using the same algorithm to perform unsupervised cluster analysis in the independent validation set (asthma $n=72$, COPD $n=42$) also resulted in five significantly different eNose-driven clusters (V clusters) (figure 3 and online supplementary figure 3E). Figure 3 shows similar positions of the clusters between training and validation sets (table 2 and online supplementary table E2). Again, all clusters were combined asthma/COPD clusters that differed regarding ethnicity ($p=0.02$), $FeNO$ ($p<0.01$), atopy ($p=0.04$), systemic eosinophilia ($p=0.05$), ICS use ($p=0.04$) and the exacerbation rate per person in the past 3 months ($p<0.01$) (online supplementary table 2E). Cluster 1V was a cluster with predominantly female patients, high BMI, low eosinophil blood counts and poor control. Type 2 inflammation was predominant in cluster 2V, with high $FeNO$. Cluster 3V consisted again of predominantly non-Caucasian patients, with the lowest registration of exacerbations in the past 3 months and no use of OCS. Cluster 4V resembled now predominantly nonallergic patients who are still frequent exacerbators. Cluster 5V consisted of fewer patients with COPD, who are younger with a normal lung function and low symptom scores. A more detailed description of the clusters obtained and similarity (based on the significant different clinical and inflammatory variables) with clusters found in the training set is presented in the online supplementary material.

TABLE 2 Training set: five eNose-driven clusters

	Cluster 1T	Cluster 2T	Cluster 3T	Cluster 4T	Cluster 5T	p-value [#]
Subjects n (%)	74 (23)	48 (15)	24 (7)	118 (37)	57 (18)	
Asthma/COPD	64.9/35.1	65.1/34.9	60.0/40.0	61.0/39.0	73.7/26.3	0.57
Age years	53.9±19.9	54.5±14.8	57.8±16.7	53.6±17.8	50.5±18.1	0.42
Female/male	83.5/16.5	39.8/61.2	55.0/45.0	57.6/49.4	57.1/42.9	0.06
BMI kg·m ⁻²	29.6±5.2	26.3±5.9	24.7±3.4	27.4±6.3	28.7±6.3	0.04
Caucasian/non-Caucasian	91.0/9.0	86.7/13.3	10.0/90.0	88.7/11.3	83.9/16.1	<0.01
Postbronchodilator FEV ₁ % predicted	79.8±29.1	78.1±26.4	68.9±26.5	78.5±25.2	83.9±22.8	0.31
F _{ENO} ppb [¶]	23.1±7.8	48.6±11.0	34.7±9.8	41.2±15.1	27.6±8.4	<0.01
	n=30	n=22	n=10	n=58	n=29	
Smoking (never/ex/current) n/n/n	31/35/8	19/25/4	8/13/3	41/58/19	26/26/5	0.57
Smoking pack-years	15.4±21.2	17.9±23.9	26.3±34.5	20.1±34.9	15.6±24.1	0.24
Atopic/non-atopic [¶]	54.7/45.3	58.1/41.9	61.1/38.9	84.0/16.0	52.5/47.5	<0.01
	n=42	n=31	n=18	n=75	n=40	
Eosinophil blood counts ×10 ⁹ cells·L ⁻¹	0.23±1.1	0.69±1.2	0.45±1.3	0.34±0.6	0.31±0.2	0.02
Neutrophil blood counts ×10 ⁹ cells·L ⁻¹	4.98±2.2	5.44±2.9	6.35±3.2	8.44±3.1	5.49±2.8	0.03
ACQ	2.87±5.5	1.80±1.1	2.36±1.4	2.17±2.8	1.73±0.9	0.71
CCQ	4.98±2.2	2.36±0.5	3.89±4.6	2.73±1.5	2.73±1.3	0.98
Exacerbations requiring treatment in past 3 months n per patient	0.22±0.4	0.25±0.5	0.04±0.2	0.43±0.3	0.14±0.3	<0.01
ICS use	70.3	83.7	55.0	66.7	78.2	0.05
OCS use	14.9	4.2	0	13.6	10.5	0.18
Food/drink consumption in 2 h prior to exhaled breath measurement	83.9	95.3	90.0	84.5	80.4	0.28

Data are presented as % or mean±SD, unless otherwise stated. Bold type represents statistical significance (p<0.05). COPD: chronic obstructive pulmonary disease; BMI: body mass index; FEV₁: forced expiratory volume in 1 s; F_{ENO}: exhaled nitric oxide fraction; ACQ: Asthma Control Questionnaire; CCQ: Clinical COPD Questionnaire; ICS: inhaled corticosteroids; OCS: oral corticosteroids. #: values obtained from analysis of variance or Chi-squared analysis between five clusters; ¶: analysis in subsets of patients.

Supervised analysis

First, a multiple linear regression was calculated in the training set to predict neutrophilic proportions measured in blood based on eNose data (PC1–4). A significant regression equation was found ($F(1,319) = 220.763$, $p < 0.001$), with $R^2 = 0.409$ (figure 4a). Participants' predicted neutrophilic proportion was equal to $0.618 + (0.093 \times PC2)$. Neutrophilic proportions increased 9.3% for each unit of PC2. Only PC2, which explained 21.3% of the variance within the dataset, was a significant predictor of neutrophilic proportions measured in blood.

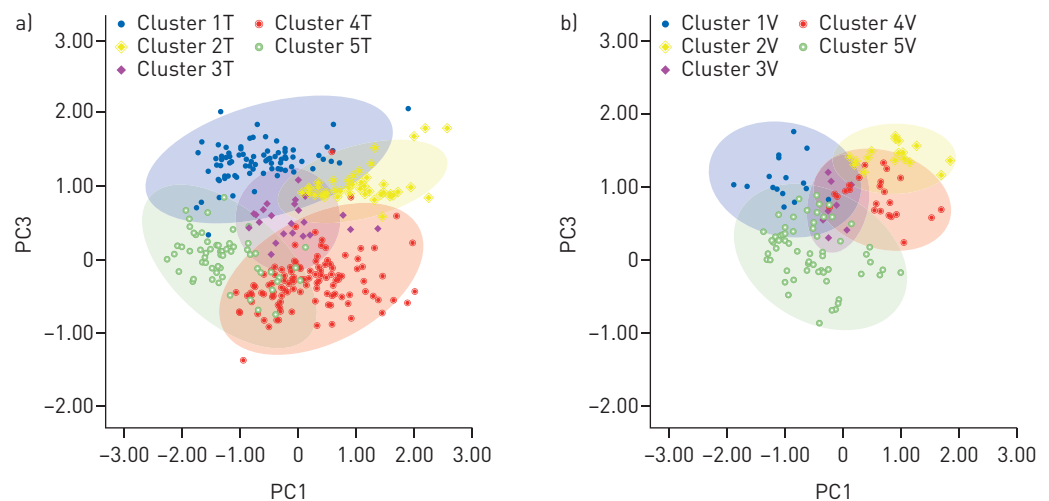


FIGURE 3 Graphical representation (principal component (PC) analysis plot) of the obtained clusters in our a) training and b) validation sets showing comparable positions of the clusters between the training and validation sets. For example, based on the significantly different clinical and inflammatory variables, cluster number 1T in the training set resembles cluster 1V of the validation set (table 2 and online supplementary table 2E), which corresponds with the position of the clusters and colours of the clusters between a) and b).

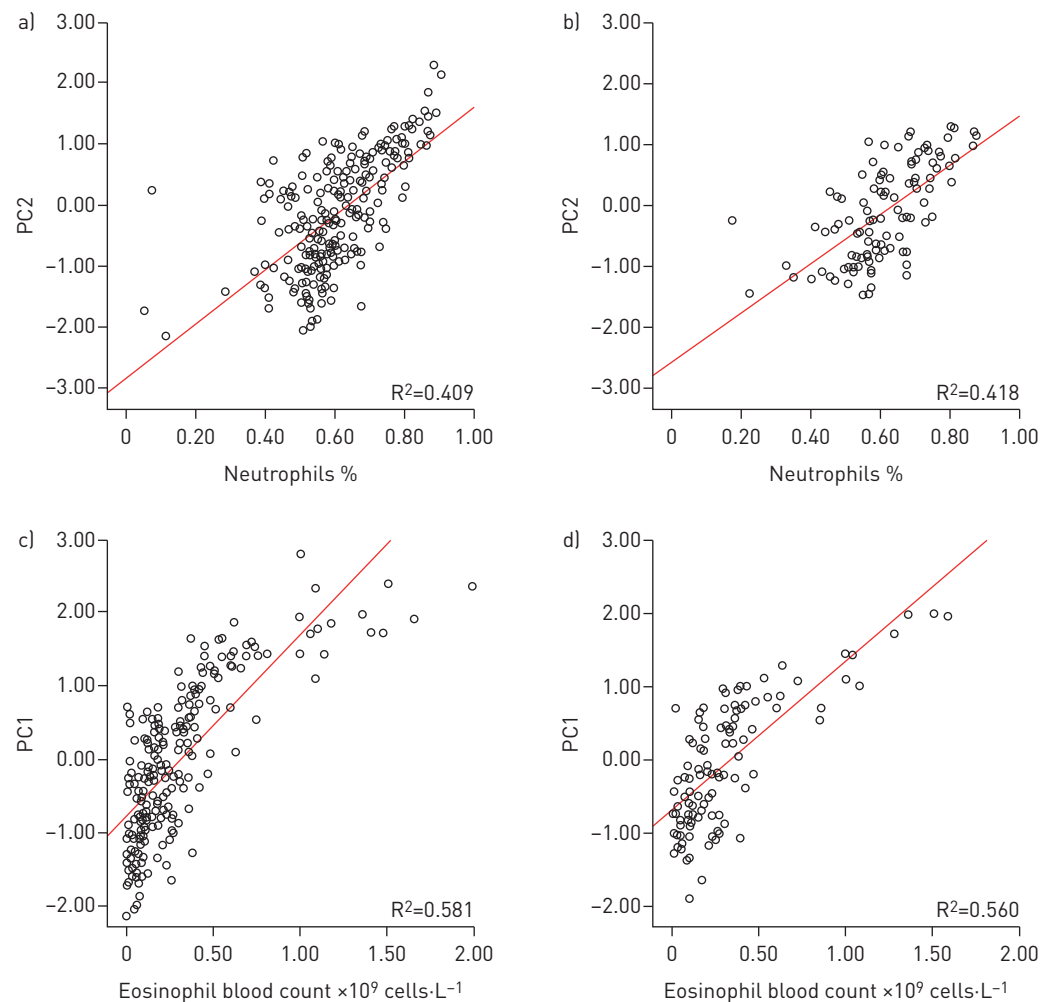


FIGURE 4 Regression plots showing the relationship between eNose (principal component (PC)2) and neutrophilic proportions a) training set, b) validation set and the relationship between eNose (PC1) and eosinophil blood counts c) training set, d) validation set.

Second, a multiple linear regression was calculated in the training set to predict eosinophil blood counts based on PC1–4. A significant regression equation was found ($F(1,319)=442.337$, $p<0.001$), with $R^2=0.581$ (figure 4c). Participants' predicted eosinophil blood counts is equal to $0.304+(0.235 \times \text{PC1})$. Eosinophilic blood counts increased 0.235×10^9 cells·L⁻¹ for each unit of PC1. PC1 that explained 42.6% of the variance within the dataset was significant in the prediction of eosinophilic blood counts.

Finally, the regression models to predict neutrophilic proportions and eosinophil blood counts obtained from the training set were confirmed in the independent validation set (figure 4b and d). More details about the results can be found in the online supplementary material.

Discussion

This study shows that phenotyping chronic airway diseases solely driven by breathomics provides combined asthma/COPD clusters that differed regarding ethnicity, systemic eosinophilia and systemic neutrophilia, $F_e\text{NO}$, BMI, atopy and exacerbation rate. This was confirmed in an independent validation set of patients. Interestingly, the eNose clusters found in a mixed sample of asthma and COPD patients resemble previously published clinical clusters in asthma or COPD patients only. Furthermore, exhaled VOCs appear to be related to systemic neutrophilic and eosinophilic blood counts among patients with chronic airway disease with evidence of a dose-dependent relationship. When considering the clinical and biological variables from the multiple regression analysis, the main independent drivers of the eNose signal in asthma and COPD appear to be systemic eosinophilia and neutrophilia, BMI, atopy, ethnicity, smoking pack-years and ICS use. These findings indicate that breathomics captures clinical/inflammatory phenotypes, regardless of the diagnosis, among patients with chronic airway disease.

This study captured clusters based on breathomics that were not predominantly driven by diagnosis but were instead driven by clinical/inflammatory characteristics. Most of these driving factors are commonly identified when attempting to phenotype asthma and/or COPD patients [25–30]. Cluster 1 found in our study contains predominantly females with a high BMI, no inflammation and high symptom scores. This cluster resembles one of the clusters found by HALDAR *et al.* [26]: “obese non-eosinophilic” asthmatics. MOORE *et al.* [27] also found a cluster of mostly older obese women with non-atopic asthma. Cluster 2 consists predominantly of males with high circulating eosinophil counts and relatively low symptom scores. This is a cluster that resembles the “inflammation predominant” phenotype found by HALDAR *et al.* [26] and is comparable with one of the clusters found by NEWBY *et al.* [25]: “late onset, eosinophilic”. The largest cluster obtained in our study, cluster 4, has much in common with the “early onset atopic” phenotype found by both NEWBY *et al.* [25] and HALDAR *et al.* [26], as patients are mainly atopic with signs of inflammation. Finally, cluster 5 is highly comparable to the “benign asthma” phenotype in which patients are well controlled without inflammation and a normal lung function. Taken together, the clusters we have obtained solely by breathomics are comparable with other clusters described in literature, which raises confidence in the robustness of these clusters.

As previously mentioned, the clusters found in our study do not represent homogenous asthma or COPD clusters, but represent combined clusters of asthma and COPD patients. RENNARD *et al.* [29] identified five COPD subgroups that differed in clinical outcomes and inflammatory biomarkers, suggesting the clusters represented clinically and biologically different subtypes of COPD. Our findings emphasise the concept that there are different phenotypes of chronic airway disease including overlapping and complementary disease entities. The latter extends the work by ROOTMENSEN *et al.* [31] and WEATHERALL *et al.* [32] who identified overlapping clinical phenotypes of obstructive airway diseases among patients with asthma or COPD by cluster analysis. Our data suggest that this also holds for inflammatory features, thereby supporting a precision medicine approach based on treatable traits rather than diagnostic labels in categorising the clinical and biological complexity of chronic airway diseases [3].

Cluster 3 obtained in this study consists mostly of non-Caucasian (90%) subjects. Apparently, ethnicity was a major driving factor in this cluster analysis, which might be explained by a complex interplay of factors including genetics, environment, culture, behaviour and socioeconomic position [33]. It has already been shown by SCHATZ *et al.* [28] that ethnicity is a variable that distinguishes different asthmatic phenotypes and that it is also linked to clinical outcomes. Notably, eosinophilic airway inflammation has recently been associated with ethnicity [34]. It will therefore be crucial to study the mechanisms that link ethnic inequalities to health, disease and phenotyping in order to individualise the management of chronic airway diseases [28].

Our study extends the results of several research groups who observed associations between eosinophilic or neutrophilic inflammatory profiles and exhaled VOCs in asthma or COPD using GC-MS [12, 13, 35]. Interestingly, in an *in vitro* study SCHLEICH *et al.* [35] identified seven VOCs that were able to discriminate between eosinophilic and neutrophilic cultures. The present study extends this to a real-life clinical setting. To our knowledge our study is the first that applied an eNose, suitable for use in a doctor’s office, to phenotype a combined sample of asthma and COPD patients based on clinical as well as inflammatory characteristics.

In addition, PLAZA *et al.* [36] aimed to investigate the capacity of eNose technology to discriminate between inflammatory phenotypes determined by induced sputum in patients with persistent asthma. In our study inflammation was determined and validated in peripheral blood, which is more and more common in daily practice, regardless of an *a priori* diagnosis of asthma or COPD. However, the association between airway (sputum) inflammatory cell counts and those in peripheral blood are at best moderate. Our group recently published a systematic review on this regarding eosinophil counts, showing a receiver operating characteristic area under the curve of 0.78 by blood eosinophils for sputum eosinophils [37]. The accuracy of blood neutrophils for sputum neutrophils is equally modest [38]. Nevertheless, novel anti-eosinophilic [8, 39] or anti-neutrophilic [40] strategies are increasingly based on measurements performed in blood, due to the difficulties of obtaining adequate sputum cell counts. Therefore, we compromised, as reaching similar, high sample sizes of patients with sputum cell counts in a large training and validation set is unrealistic.

The strength of our study seems to be the use of an independent validation set that confirmed the results from the training set. However, the match between the results obtained from the training and validation sets was not complete. Peripheral blood neutrophil counts were significantly different between clusters in the training set, but not in the validation set. This could be explained by differences in baseline characteristics between the training and validation sets, in which the neutrophil blood counts were higher in the training set. In addition, BMI as driving factor between clusters was not confirmed by the validation

set. Furthermore, atopy and F_{eNO} were merely obtained in a subgroup of patients, predominantly asthmatics, hampering the generalisability of these results in patients with chronic airway disease.

This study purposely used a comprehensive analysis plan. A potential limitation to this strategy is that it was based on a cross-sectional design with collection of a single blood sample as part of routine clinical characterisation within a 3-month window prior to the eNose measurement. Some recent studies have shown a degree of variability in blood inflammatory cell counts when collected at different time points [41–44]. It might be possible that this temporal variability could confound interpretation of inflammatory cell pattern heterogeneity. However, KATZ *et al.* [45] showed that a single blood eosinophil measurement in severe eosinophilic asthma patients predicted the average of subsequent measurements in 85% of cases. Using an average of multiple measurements only marginally increased the sensitivity.

What could be the clinical relevance of our findings? An ideal biomarker is minimally invasive, easy to collect, reliable, inexpensive and can be used to identify either a clinical phenotype or a treatment response phenotype, to measure changes in disease activity or to confirm a diagnosis [45]. This study used the SpiroNose to perform exhaled breath analysis as part of routine assessment at the point of care. This real-time approach of breath sampling has the advantage of increasing VOC concentrations and stability of the total exhaled breath sample compared to indirect VOC sampling using collecting bags [17, 46]. Direct breath sampling is a very appealing option, as it is not only practical for clinical usage, but it also avoids errors during collection and storage [17]. Given this clinically applicable technology, the present data may qualify breath assessment as a real-time tool in the doctor's office for broadly phenotyping patients with chronic airway disease. This meets the demands of modern medicine, in which management decisions are taken based on the patient's individual phenotype rather than a classical diagnosis [3, 47]. Novel anti-eosinophilic [8, 39, 48] or anti-neutrophilic [40] strategies demand rapid inflammatory phenotyping, which appears to be accessible using exhaled air.

In conclusion, phenotyping of a combined sample of asthma and COPD patients driven by eNose provides combined clusters of chronic airway disease that are not determined by diagnosis, but rather by clinical/inflammatory characteristics. These clusters resemble published clusters based on clinical characteristics and are related to systemic neutrophilic and eosinophilic inflammation. Our findings support the concept of a label-free approach based on treatable traits to understand the condition of individual patients. The present data suggest that molecular profiling of exhaled air allows meaningful phenotyping of chronic airway disease, regardless of the diagnosis asthma or COPD, and may therefore facilitate personalised anti-inflammatory strategies.

References

- 1 Reddel HK, Bateman ED, Becker A, *et al.* A summary of the new GINA strategy: a roadmap to asthma control. *Eur Respir J* 2015; 46: 622–639.
- 2 Vogelmeier CF, Criner GJ, Martinez FJ, *et al.* Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary. *Eur Respir J* 2017; 49: 1700214.
- 3 Agustí A, Bel E, Thomas M, *et al.* Treatable traits: toward precision medicine of chronic airway diseases. *Eur Respir J* 2016; 47: 410–419.
- 4 Desai D, Brightling C. Cytokines and cytokine-specific therapy in asthma. *Adv Clin Chem* 2012; 57: 57–97.
- 5 Petsky HL, Cates CJ, Lasserson TJ, *et al.* A systematic review and meta-analysis: tailoring asthma treatment on eosinophilic markers (exhaled nitric oxide or sputum eosinophils). *Thorax* 2012; 67: 199–208.
- 6 Gonem S, Raj V, Wardlaw AJ, *et al.* Phenotyping airways disease: an A to E approach. *Clin Exp Allergy* 2012; 42: 1664–1683.
- 7 Siva R, Green RH, Brightling CE, *et al.* Eosinophilic airway inflammation and exacerbations of COPD: a randomised controlled trial. *Eur Respir J* 2007; 29: 906–913.
- 8 Cabon Y, Molinari N, Marin G, *et al.* Comparison of anti-interleukin-5 therapies in patients with severe asthma: global and indirect meta-analyses of randomized placebo-controlled trials. *Clin Exp Allergy* 2017; 47: 129–138.
- 9 Vermeersch K, Gabrovská M, Deslypere G, *et al.* The Belgian trial with azithromycin for acute COPD exacerbations requiring hospitalization: an investigator-initiated study protocol for a multicenter, randomized, double-blind, placebo-controlled trial. *Int J Chron Obstruct Pulmon Dis* 2016; 11: 687–696.
- 10 Pascoe S, Locantore N, Dransfield MT, *et al.* Blood eosinophil counts, exacerbations, and response to the addition of inhaled fluticasone furoate to vilanterol in patients with chronic obstructive pulmonary disease: a secondary analysis of data from two parallel randomised controlled trials. *Lancet Respir Med* 2015; 3: 435–442.
- 11 Bos LD, Sterk PJ, Fowler SJ. Breathomics in the setting of asthma and chronic obstructive pulmonary disease. *J Allergy Clin Immunol* 2016; 138: 970–976.
- 12 Ibrahim B, Basanta M, Cadden P, *et al.* Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 2011; 66: 804–809.
- 13 Fens N, de Nijs SB, Peters S, *et al.* Exhaled air molecular profiling in relation to inflammatory subtype and activity in COPD. *Eur Respir J* 2011; 38: 1301–1309.
- 14 Wilson AD. Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. *Metabolites* 2015; 5: 140–163.
- 15 Boots AW, Bos LD, van der Schee MP, *et al.* Exhaled molecular fingerprinting in diagnosis and monitoring: validating volatile promises. *Trends Mol Med* 2015; 21: 633–644.

- 16 Fens N, van der Schee MP, Brinkman P, *et al.* Exhaled breath analysis by electronic nose in airways disease. Established issues and key questions. *Clin Exp Allergy* 2013; 43: 705–715.
- 17 de Vries R, Brinkman P, van der Schee MP, *et al.* Integration of electronic nose technology with spirometry: validation of a new approach for exhaled breath analysis. *J Breath Res* 2015; 9: 046001.
- 18 Bateman ED, Hurd SS, Barnes PJ, *et al.* Global strategy for asthma management and prevention: GINA executive summary. *Eur Respir J* 2008; 31: 143–178.
- 19 Juniper EF, O'Byrne PM, Guyatt GH, *et al.* Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999; 14: 902–907.
- 20 Reda AA, Kotz D, Kocks JW, *et al.* Reliability and validity of the clinical COPD questionnaire and chronic respiratory questionnaire. *Respir Med* 2010; 104: 1675–1682.
- 21 Broadhurst DI, Kell DB. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* 2006; 2: 171–196.
- 22 Yeomans KA, Golder PA. The Guttman–Kaiser criterion as a predictor of the number of common factors. *J Royal Stat Soc Ser D* 1982; 31: 221–229.
- 23 Brinkman P, van de Pol MA, Gerritsen MG, *et al.* Exhaled breath profiles in the monitoring of loss of control and clinical recovery in asthma. *Clin Exp Allergy* 2017; 47: 1159–1169.
- 24 Clarke KR, Somerfield PJ, Gorley RN. Testing of null hypotheses in exploratory community analyses: similarity profiles and biota–environment linkage. *J Exp Mar Biol Ecol* 2008; 366: 56–69.
- 25 Newby C, Heaney LG, Menzies-Gow A, *et al.* Statistical cluster analysis of the British Thoracic Society Severe Refractory Asthma Registry: clinical outcomes and phenotype stability. *PLoS One* 2014; 9: e102987.
- 26 Halder P, Pavord ID, Shaw DE, *et al.* Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008; 178: 218–224.
- 27 Moore WC, Meyers DA, Wenzel SE, *et al.* Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010; 181: 315–323.
- 28 Schatz M, Hsu JW, Zeiger RS, *et al.* Phenotypes determined by cluster analysis in severe or difficult-to-treat asthma. *J Allergy Clin Immunol* 2014; 133: 1549–1556.
- 29 Rennard SI, Locantore N, Delafont B, *et al.* Identification of five chronic obstructive pulmonary disease subgroups with different prognoses in the ECLIPSE cohort using cluster analysis. *Ann Am Thorac Soc* 2015; 12: 303–312.
- 30 Górka K, Paplińska-Goryca M, Nejman-Gryz P, *et al.* Eosinophilic and neutrophilic airway inflammation in the phenotyping of mild-to-moderate asthma and chronic obstructive pulmonary disease. *COPD* 2017; 14: 181–189.
- 31 Rootmensen G, van Keimpema A, Zwinderman A, *et al.* Clinical phenotypes of obstructive airway diseases in an outpatient population. *J Asthma* 2016; 53: 1026–1032.
- 32 Weatherall M, Travers J, Shirtcliffe PM, *et al.* Distinct clinical phenotypes of airways disease defined by cluster analysis. *Eur Respir J* 2009; 34: 812–818.
- 33 Stronks K, Snijder MB, Peters RJ, *et al.* Unravelling the impact of ethnicity on health in Europe: the HELIUS study. *BMC Public Health* 2013; 13: 402.
- 34 Nyenhuis SM, Krishnan JA, Berry A, *et al.* Race is associated with differences in airway inflammation in patients with asthma. *J Allergy Clin Immunol* 2017; 140: 257–265.
- 35 Schleich FN, Dallinga JW, Henket M, *et al.* Volatile organic compounds discriminate between eosinophilic and neutrophilic inflammation *in vitro*. *J Breath Res* 2016; 10: 016006.
- 36 Plaza V, Crespo A, Giner J, *et al.* Inflammatory asthma phenotype discrimination using an electronic nose breath analyzer. *J Investig Allergol Clin Immunol* 2015; 25: 431–437.
- 37 Korevaar DA, Westerhof GA, Wang J, *et al.* Diagnostic accuracy of minimally invasive markers for detection of airway eosinophilia in asthma: a systematic review and meta-analysis. *Lancet Respir Med* 2015; 3: 290–300.
- 38 Hastie AT, Moore WC, Li H, *et al.* Biomarker surrogates do not accurately predict sputum eosinophil and neutrophil percentages in asthmatic subjects. *J Allergy Clin Immunol* 2013; 132: 72–80.
- 39 Wang FP, Liu T, Lan Z, *et al.* Efficacy and safety of anti-interleukin-5 therapy in patients with asthma: a systematic review and meta-analysis. *PLoS One* 2016; 11: e0166833.
- 40 Bruijnzeel PL, Uddin M, Koenderman L. Targeting neutrophilic inflammation in severe neutrophilic asthma: can we target the disease-relevant neutrophil phenotype? *J Leukoc Biol* 2015; 98: 549–556.
- 41 Acland JD, Gould AH. Normal variation in the count of circulating eosinophils in man. *J Physiol* 1956; 133: 456–466.
- 42 Spector SL, Tan RA. Is a single blood eosinophil count a reliable marker for “eosinophilic asthma?” *J Asthma* 2012; 49: 807–810.
- 43 Szeffler SJ, Wenzel S, Brown R, *et al.* Asthma outcomes: biomarkers. *J Allergy Clin Immunol* 2012; 129: Suppl., S9–S23.
- 44 Brunck ME, Andersen SB, Timmins NE, *et al.* Absolute counting of neutrophils in whole blood using flow cytometry. *Cytometry A* 2014; 85: 1057–1064.
- 45 Katz LE, Gleich GJ, Hartley BF, *et al.* Blood eosinophil count is a useful biomarker to identify patients with severe eosinophilic asthma. *Ann Am Thorac Soc* 2014; 11: 531–536.
- 46 Montuschi P, Mores N, Trovè A, *et al.* The electronic nose in respiratory medicine. *Respiration* 2013; 85: 72–84.
- 47 Pavord ID, Wardlaw AJ. The A to E of airway disease. *Clin Exp Allergy* 2010; 40: 62–67.
- 48 George L, Brightling CE. Eosinophilic airway inflammation: role in asthma and chronic obstructive pulmonary disease. *Ther Adv Chronic Dis* 2016; 7: 34–51.