Original Article

Study of 5 Volatile Organic Compounds in Exhaled Breath in Chronic Obstructive Pulmonary Disease

José Javier Jareño-Esteban, a M. Ángeles Muñoz-Lucas, b,c Óscar Gómez-Martín, c Sergio Utrilla-Trigo, d Carlos Gutiérrez-Ortega, e Antonio Aguilar-Ros, f Luis Collado-Yurrita, g Luis Miguel Callol-Sánchez h

a Servicio de Neumología, Hospital Central de la Defensa, Universidad de Alcalá de Henares, IMIDEF, Alcalá de Henares, Madrid, Spain
b Instituto Mixto de Investigación Biosanitaria de la Defensa (IMIDEF), Hospital Central de la Defensa, Madrid, Spain
c Instituto Mixto de Investigación Biosanitaria de la Defensa (IMIDEF), Madrid, Spain
d Universidad Complutense de Madrid, Madrid, Spain
e Unidad de Apoyo a la Investigación, Hospital Central de la Defensa, Madrid, Spain
f Dpto. CC Farmacéuticas, Facultad de Farmacia, Universidad San Pablo CEU, Madrid, Spain
g Dpto. de Medicina, Universidad Complutense de Madrid, Madrid, Spain
h Dpto. de Medicina, Facultad de Medicina, Universidad de Alcalá, Instituto Mixto de Investigación Biosanitaria de la Defensa (IMIDEF), Alcalá de Henares, Madrid, Spain

ARTICLE INFO

Article history:
Received 7 March 2016
Accepted 7 September 2016
Available online 12 April 2017

Keywords:
Chronic obstructive pulmonary disease
Volatile organic compounds
Exhaled breath
Oxidative stress

ABSTRACT

Introduction: A major risk factor for chronic obstructive pulmonary disease (COPD) is tobacco smoke, which generates oxidative stress in airways, resulting in the production of volatile organic compounds (VOCs). The purpose of this study was to identify VOCs in exhaled breath and to determine their possible use as disease biomarkers.

Method: Exhaled breath from 100 healthy volunteers, divided into 3 groups (never smokers, former smokers and active smokers) and exhaled breath from 57 COPD patients were analyzed. Samples were collected using BioVOC® devices and transferred to universal desorption tubes. Compounds were analyzed by thermal desorption, gas chromatography and mass spectrometry. VOCs analyzed were linear aldehydes and carboxylic acids.

Results: The COPD group and healthy controls (never smokers and former smokers) showed statistically significant differences in hexanal concentrations, and never smokers and the COPD group showed statistically significant differences in nonanal concentrations.

Conclusions: Hexanal discriminates between COPD patients and healthy non-smoking controls. Nonanal discriminates between smokers and former smokers (with and without COPD) and never smokers.

© 2016 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

Estudio de 5 compuestos orgánicos volátiles en aire exhalado en la enfermedad pulmonar obstructiva crónica

RÉSUMÉ

Introducción: Un factor de riesgo importante para el desarrollo de la enfermedad pulmonar obstructiva crónica (EPOC) es el humo del tabaco, que genera estrés oxidativo en las vías respiratorias, dando lugar a la producción de compuestos orgánicos volátiles (VOC). El objetivo del trabajo es su identificación en el aire exhalado y su posible utilidad como biomarcadores de la enfermedad.

Método: Se analizó el aire exhalado de 100 voluntarios sanos, clasificados en 3 grupos (no fumadores, exfumadores y fumadores activos) y un grupo de 57 pacientes con EPOC. La muestra de aire exhalado se recogió mediante BioVOC® y se traspuso a tubos de desorción para su posterior análisis por cromatografía de gases y espectrometría de masas. Los VOC analizados fueron aldehídos lineales y ácidos carboxílicos.

© 2016 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.

 Palabras clave:
Enfermedad pulmonar obstructiva crónica
Compuestos orgánicos volátiles
Aire exhalado
Estrés oxidativo


 Corresponding author.
E-mail address: mmunozlucas@yahoo.es (M.Á. Muñoz-Lucas).

1579-2129 © 2016 SEPAR. Published by Elsevier España, S.L.U. All rights reserved.
Introduction

Chronic obstructive pulmonary disease (COPD) is defined as a respiratory disease characterized by chronic, progressive airflow limitation that is not fully reversible. The main symptoms are dyspnea and cough, sometimes accompanied by expectoration. COPD patients have exacerbations that vary in severity depending on comorbidities.1

Smoking is the most important risk factor in the development of COPD, and evidence shows that the risk is proportional to the accumulated consumption of cigarettes.2 In Spain, 7.6% of male and 5.5% of female non-smokers have COPD,3 but this figure rises to 39.5% and 15.4% among men and women, respectively, who have smoked for more than 10 years. Other factors, such as alpha-1 antitrypsin deficiency or pneumonia in childhood, also play a role.4

According to the World Health Organization (WHO), 2.9 million individuals a year die from COPD, and it is estimated that by 2030, it will be the third cause of death worldwide.5

In the EPI-SCAN epidemiological study of COPD in Spain,6 73% of confirmed cases did not have a previous diagnosis of COPD, revealing a high degree of underdiagnosis. The standard technique is spirometry, but this method carries a risk of underdiagnosis in the early stages and overdiagnosis in the advanced stages. Other useful methods are available, such as bronchoalveolar lavage or open lung biopsy, but these are too invasive for routine use.7

Clinicians involved in diagnosing COPD are continually on the lookout for techniques and parameters that will help them in their decision-making. Biomarkers are biological parameters that provide information on the normal or disease status of an individual or a population.9 The search for biomarkers to characterize COPD is ongoing, and possible candidates have recently been investigated in sputum,8 bronchoalveolar lavage,9 and exhaled air.10,11

Exhaled air contains a multitude of volatile organic compounds (VOCs), some of which can be identified as biomarkers that may be of use in the characterization of COPD. The analytical procedure is non-invasive and rapid, and could complement spirometry in both diagnosis and follow-up of this entity. However, this technique is susceptible to contamination of samples by multiple environmental compounds, and this needs to be taken into account when interpreting results.

Tobacco smoke contains over 2000 compounds and a great quantity of free radicals and reactive oxygen and nitrogen species, which increase oxidative stress and pulmonary inflammation.12 Increased oxidative stress causes lipid peroxidation. The damage generated by this chain reaction produces a large amount of VOCs, including alkanes, aldehydes, and carboxylic acids, which may be excreted by the airways. The presence of these compounds in exhaled air may be a sign of oxidative stress in the airways and the lungs. These VOCs meet the definition of inflammatory biomarkers mentioned above.

A series of benchmark studies on VOCs as biomarkers for COPD in COPD patients compared to clinically healthy controls have been published.14–18 The studies are similar, but no clear conclusions can be reached as the results vary widely.

The aim of our study was to determine if significant differences really exist between certain VOCs found in the exhaled air of COPD patients compared to healthy controls, and if these substances could be considered as disease biomarkers.

Patients and Methods

This was a case–control study with consecutive non-probability sampling. A total of 157 volunteers were selected among the employees and patients of the Hospital Central de la Defensa “Gomez Ulla” and the General Air Force Headquarters between October 2014 and December 2015. Two study groups were established, one consisting of 57 clinically stable COPD patients, and the other of 100 healthy controls (never smokers, former smokers, and active smokers).

The inclusion criteria for the 2 groups included consent to participate in the study, age over 40 years, and a smoking status of never, former or active (according to WHO criteria). All individuals completed a questionnaire and an additional clinical examination, including a flow-volume loop. Patients in the COPD group underwent standard tests for the diagnosis and follow-up of their disease, including chest radiograph, flow-volume loop, bronchodilator testing, etc. COPD severity was classified according to the GOLD scale.

Exclusion criteria consisted of any other current or previous lung or tumor disease of any organ or system, or refusal to participate in the study. No gender-based restrictions were applied.

None of the participants were exposed to any special occupational environmental conditions.

Subjects were informed about the aims, risks and benefits, and planned tests and techniques used for conducting the study. They were given written information before signing the informed consent form. All study data were handled in accordance with the provisions of Organic Law 15/1999 on the Protection of Personal Data, 13 December 1999, and Act 41/2002, 14 November 2002, regulating the autonomy of patients and their rights and obligations in relation to clinical information and documentation. The study protocol was approved by the Ethics and Clinical Research Committee of Hospital Central de la Defensa “Gomez Ulla”.

This was a targeted chromatographic study with previous selection of the compounds to be studied. VOCs were selected as follows: (1) contaminants derived from the environment or from tobacco smoke were excluded; (2) their metabolic origin had to be known for them to meet criteria for use as biomarkers.

Thus, of more than 250 VOCs in exhaled air described in the consulted literature, only 50 were preselected, on the basis of frequency. Some were endogenous compounds derived from lipid peroxidation. Others were environmental contaminants, and, lastly, another group were compounds of undetermined origin. These latter 2 groups were ruled out, as the origin of the VOCs had to be known in order to select compounds that were of real use as possible biomarkers.

The final selection produced 5 VOCs that met the required criteria: 3 linear, hexanal, heptanal and nonanal aldehydes (known metabolites of omega 3, omega 6 and omega 9 fatty acid lipid peroxidation); and 2 carboxylic acids, propanoic acid and nonanoic acid (also metabolites of fatty acid lipid peroxidation).11,12

After individuals had rested for 1 h, with no oral intake or smoking, the exhaled air sample was collected in BioVOC® breath samples. These easy-to-manage devices do not generate resistance
to the passage of air and can be used for the analysis of a fixed volume. The exhaled air sample was obtained with a forced expiration maneuver. The air corresponding to the anatomical dead space passes through the BioVOC” sampler, and the most representative possible fraction of alveolar air is collected in the chamber. The maneuver was repeated 3 times for preconcentration of the compounds. The collected air was transferred to a thermal desorption tube with capacity for absorbing compounds from C2 to C20.

A sample of ambient air was simultaneously obtained from the examination room and processed in the same way as the exhaled air in order to compare VOC levels from ambient and exhaled air, and to determine endogenic or exogenic origin.

All samples were stored at room temperature. The manufacturer states that samples are stable for up to 6 days, but all were analyzed within a maximum of 24 h after sampling to minimize any possible cross-reactions between the compounds, given the high concentration of water in the exhaled air.

The sampling method is described in a previous paper published by our group.12

Two types of results were obtained depending on the limit of detection (LOD): undetected compounds, with a value below the LOD, and detected compounds, with a value above the LOD.

A compound was classified as detected after 3 stages:

1. The area of the quantifier ion of the mass spectrum was integrated for each VOC.
2. The area of each compound was normalized according to the area of the 207 ion (corresponding to hexamethyldicyclosiloxane). Inorganic compound, obtained from the desorption tubes that does not interfere with the VOCs, was used as an internal reference compound. Compounds were normalized by measuring the area under the curve using the following formula:

\[
\frac{(A_{bc,VOC})}{(A_{bc,Siloxane})} \times 100
\]

where \( A_{bc} \) is the area under the curve of each compound.
3. A second normalization was conducted between now normalized exhaled air and ambient air. A compound was assumed to be detected when the ratio between the value in exhaled air/ambient air was greater than 1, or when it was detected in exhaled air and not in ambient air.

**Variables**

A diagnosis of COPD was taken as an independent variable. Qualitative dependent variables were the different VOCs, classified as undetected and detected. Demographic variables were sex and age.

VOC distributions were verified using the Kolmogorov–Smirnov test, and median and interquartile range (IQR) were used as measures of central tendency and dispersion. Dependent variables were expressed as relative frequencies.

The association between independent and dependent categorical variables was confirmed by odds ratio, based on logistic regression and the corresponding 95% confidence interval.

The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of VOCs were calculated to evaluate their use as a diagnostic test for COPD.

A P-value of <.05 was considered statistically significant, and the statistical package used was SPSS® version 20.

**Results**

Table 1 summarizes the demographic characteristics of the study population. The control group comprised 100 healthy volunteers, with a mean age of 49 (9) years, and the COPD group comprised 57 patients, mean age 73.6 (9.5) years. Distribution by sex was 49% men and 51% women in the control group, and 89% men and 11% women in the COPD group. Thirty-three percent of the control group were smokers, 28% former smokers, and 39% never smokers, with a mean pack-year index (PYI) of 29 (20). In the COPD group, 14% were smokers, 86% were former smokers, PYI 74.4 (39.2). Fourteen percent of smokers were GOLD 1, 31.6% GOLD 2, 24.6% GOLD 3, and 29.8% GOLD 4.

The 2 categories differentiated by the values obtained were (a) detected and (b) undetected, according to the criteria defined in the section “Patients and methods”. Frequency distribution, which was non-normal, is shown in percentages in Fig. 1.

**Table 1** Study Group Demographics.

<table>
<thead>
<tr>
<th></th>
<th>Control (n=100)</th>
<th>COPD (n=57)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>48.9 (9.4)</td>
<td>73.6 (9.5)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>49 (49%)</td>
<td>51 (89%)</td>
</tr>
<tr>
<td>Women</td>
<td>51 (51%)</td>
<td>6 (11%)</td>
</tr>
<tr>
<td>Smokers</td>
<td>33 (33%)</td>
<td>8 (14%)</td>
</tr>
<tr>
<td>Former smokers</td>
<td>28 (28%)</td>
<td>49 (86%)</td>
</tr>
<tr>
<td>Never smokers</td>
<td>39 (39%)</td>
<td></td>
</tr>
<tr>
<td>Smoking habit (PYI), mean (SD)</td>
<td>29 (20)</td>
<td>74.4 (39.2)</td>
</tr>
</tbody>
</table>

PYI: pack-year index; SD: standard deviation.

**Fig. 1.** Detection of study VOCs by frequency.

**Table 2** Relationship Between the Presence of Markers and the Presence of COPD Compared to the Control Group.

<table>
<thead>
<tr>
<th>VOC</th>
<th>Control</th>
<th>COPD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hexanal</td>
<td>1</td>
<td>1.97 (1.09–3.56)</td>
<td>.023</td>
</tr>
<tr>
<td>Heptanal</td>
<td>1</td>
<td>0.70 (0.39–1.26)</td>
<td>.223</td>
</tr>
<tr>
<td>Nonanal</td>
<td>1</td>
<td>1.46 (0.96–2.22)</td>
<td>.08</td>
</tr>
<tr>
<td>Propanoic acid</td>
<td>1</td>
<td>1.05 (0.77–1.42)</td>
<td>.773</td>
</tr>
<tr>
<td>Nonanoic acid</td>
<td>1</td>
<td>1.42 (0.74–2.75)</td>
<td>.291</td>
</tr>
</tbody>
</table>

OR and 95% confidence interval based on logistic regression.
When the COPD group was compared with the never smoker control subgroup (Table 3), significant differences were again detected in hexanal (P=.015), with a 3.08-fold probability of detecting this VOC in COPD patients than in never smokers (95% CI: 1.13–8.40) (sensitivity 31.58%, specificity 89.74%, PPV 81.82% and NPV 47.30%). Significant differences in nonanal (P=.008) were also observed between never smoker healthy controls and the COPD group (smokers and former smokers pooled), with a 2.44-fold probability of detecting nonanal in COPD patients (95% CI: 1.17–5.08).

When the COPD group was compared with the active smoker control subgroup (Table 3 and Fig. 2), no statistical significance was found in the VOCs under study. However, when the COPD group was compared with the former smoker control subgroup (Table 3), the presence of hexanal was found to be statistically significant. The probability of detecting hexanal in the COPD group is 4.42 times higher than in the former smoker control subgroup (95% CI: 1.10–17.73) (sensitivity 31.57%, specificity 92.85%, PPV 90% and NPV 40%).

**Discussion**

The main aim of this study was to determine differences between the 5 selected VOCs (hexanal, heptanal, nonanal, propanoic acid and nonanoic acid) in a COPD group and a presumably healthy control group.

Statistically significant differences were found for hexanal, with a 1.94-fold probability of detecting this VOC in patients with COPD. This aldehyde is an end metabolite of lipid peroxidation of phospholipids that form the cell membranes. Its detection in COPD patients and not in the control group suggests that oxidative stress is greater in the lower respiratory tract of the COPD patients. For a better understanding of the differences between the control group and COPD group, the controls were subdivided into “never smoker control subgroup”, “former smoker control subgroup”, and “smoker control subgroup”, and each was compared separately with the COPD group.

In the first comparison between the never smoker control subgroup and the COPD group, hexanal concentrations were significantly different, with a 3.08-fold greater probability of detection in the COPD group. Likewise, differences in nonanal levels were also significant (P=.008). This latter finding is in line with results obtained in a previous paper, in which nonanal was associated with smoking.

A similar pattern emerged for hexanal when the COPD group was compared with the former smoker control subgroup (P=.013).

In contrast, no significant differences were found between the COPD group and the active smoker control subgroup. Our interpretation is that inflammation due to active smoking might mask COPD-specific inflammation, and that the production of VOCs due to the disease becomes of secondary importance.

For the same reason, it is logical to suppose that the significant differences in hexanal between the COPD group and the former smoker control groups are due, in general, to patients quitting smoking at the time of receiving their COPD diagnosis, with the subsequent resolution of the direct inflammatory effect of tobacco.

This finding confounds the evaluation of hexanal as a discriminant biomarker for COPD and rules out its use in the detection of COPD in asymptomatic smokers, one of the targets of this
study. However, the low rate of COPD among active smokers (8 of the 57) means that further, larger studies are needed in this subpopulation.

PPV and specificity of hexanal in the COPD group and the never-smoker and former smoker subgroups are high. In the latter case, this finding may be a useful screening tool in clinical practice, although at present the cost/benefit ratio is high compared to standard spirometry.

The wide variety of analytical methods and instruments available has produced disparate results. Corradi et al. for example, use high-resolution liquid chromatography and mass spectrometry, while Basanta et al. use gas chromatography and time-of-flight mass spectrometry. Both techniques, though still accepted, have certain features that make them unsuitable in this context. The first requires the gases to be dissolved in a liquid medium, and the cost-effectiveness ratio of the second is very unfavorable. In contrast, our technique—gas chromatography and mass spectrometry (GC/MS)—is used by most authors and offers good, cost-effective results.\(^\text{12,18,20}\)

The large range of sampling equipment (BioVOC\textsuperscript{®}, SPME, Tedlar bags, desorption tubes, etc.) also leads to widely varying results. In our study, we used a standard method that is useful for detecting a wide range of compounds (from C2 to C20), favoring the sustainability of the study.

Taking these considerations into account, our results would be comparable with those reported by Phillips et al.,\(^\text{18}\) who used the same methodology and instrumentation. These authors, however, only found differences in isoprene, a compound associated with the synthesis of cholesterol, which is not a product of lipid peroxidation. It is considered a confounding factor, and for this reason was excluded from our study.

Researchers must be wary when selecting compounds as proposed biomarkers. Van Berkel et al.\(^\text{17}\) found differences between COPD patients and healthy controls in a panel of 13 VOCs, and when results were restricted to 6 of these, sensitivity of 100\% and specificity of 81\% was achieved. In our opinion, branched hydrocarbons, which are all of uncertain origin, together with environmental contaminants and contaminants originating in tobacco smoke are unsuitable candidates for evaluation as biomarkers. Kischkel et al.,\(^\text{20}\) like us, found that the origin of VOCs used in predictive models must be clearly defined. The fact that the presence of a compound is statistically significant does not make it suitable for inclusion as a biomarker. The identification of biomarkers must be well founded if data normalization and processing are to provide valuable clinical data.

Most COPD studies use mathematical models with several compounds, but we only obtained 1 biomarker that could differentiate between COPD patients and healthy non-active smoker subcontrols. In contrast, the model proposed by Basanta et al.\(^\text{15}\) notably increases sensitivity (90\%) when only active smokers, whether COPD or control, are selected from the study groups. This increased sensitivity occurs because the statistical power is reduced when the population size is considerably reduced, i.e., from 39 to 20 active smokers with COPD, and from 32 control group smokers to only 6. These results, which at first sight appear interesting, have a significant statistical bias in terms of sample size and subgroup division.

To sum up, hexanal discriminates between COPD patients and healthy never smoker and former smoker controls. Nonanal discriminates between smokers and former smokers (with or without COPD) and never smoker controls. It is essential to reach a consensus on the methodology and selection of compounds proposed as biomarkers.

**Funding**

This study was funded by the Instituto de Salud Carlos III (PI07/1116), SEPAR 2013 (Registry no. 135).

**Conflict of Interests**

The authors state they have no potential conflict of interests with any of the companies whose products or services may have been discussed in this article.

**Acknowledgements**

We thank the General Health Inspection Unit (IGESAN) for the use of their facilities.

We also thank the Hospital Central de la Defensa “Gómez Ulla” for providing patients and volunteers for participation in this study, and the study volunteers from the General Air Force Headquarters for their collaboration.

**References**

