Original Article

Volatile Organic Compounds in Exhaled Breath in a Healthy Population: Effect of Tobacco Smoking

José Javier Jareño-Esteban, M. Ángeles Muñoz-Lucas, Belén Carrillo-Aranda, José Ángel Maldonado-Sanz, Ignacio de Granda-Orive, Antonio Aguilar-Ros, Concepción Civera-Tejuca, Carlos Gutiérrez-Ortega, Luis Miguel Callol-Sánchez, on behalf of the Study Group

**A R T I C L E   I N F O**

Article history:
Received 1 January 2013
Accepted 9 April 2013
Available online 6 October 2013

Keywords:
Volatile organic compounds
Tobacco smoking
Oxidative stress

**A B S T R A C T**

Introduction: Tobacco smoke is a source of free radicals and reactive oxygen and nitrogen species, which are the main causes of oxidative stress. The analysis of volatile organic compounds (VOC) in exhaled breath is an indirect method of measuring the level of oxidative stress that occurs in the airways caused by tobacco consumption. The aim of this study was to determine whether smoking influences the production of VOC, in a clinically healthy population.

Methods: Exhaled breath from 89 healthy volunteers, divided into three groups (non-smokers, ex-smokers and smokers), was analyzed. Samples were collected using Bio-VOC® devices and transferred to universal desorption tubes. Chemical compounds were analyzed by thermal desorption, gas chromatography and mass spectrometry. We analyzed hexanal, heptanal, octanal, nonanal, nonanoic acid and propanoic acid, and all were identified by retention time and mass spectra referenced in the NIST 08 mass spectral library; confirmation was carried out using reference standards of the pure chemical compound.

Results: These VOC were found in very low concentrations. Only nonanal showed significant quantitative and qualitative statistical differences among the study groups. Nonanal concentration is dependent on smoking, but is independent of the amount of tobacco consumed, age and gender.

Conclusions: Nonanal in exhaled breath is associated with tobacco consumption, current or previous. Nonanal is a sub-product of the destruction of the cell membrane, and its finding may be indicative of cell damage in smokers. This result appears in many farmers who smoke.

© 2013 SEPAR. Published by Elsevier España, S.L. All rights reserved.


**R E S U M E N**

Introducción: El humo del tabaco es una fuente de radicales libres y especies reactivas de oxígeno y de nitrógeno, principales causantes de estrés oxidativo. El análisis de compuestos orgánicos volátiles (VOC)
en aire exhalado es un método indirecto de medir el nivel de estrés oxidativo que se produce en las vías aéreas. El objetivo de este trabajo es conocer la influencia del tabaco en la producción de VOC en una población clínicamente sana.

Métodos: Se analizó el aire exhalado de 89 voluntarios sanos, clasificados en 3 grupos: no fumadores, exfumadores y fumadores activos. La muestra de aire exhalado se recogió mediante Bio-VOC®, y se traspasó a tubos de desorción. La técnica analítica utilizada fue: desorción térmica, cromatografía de gases y espectrometría de masas. Los VOC analizados fueron hexanal, heptanal, octanal, nonanal, ácido propanónico y ácido nonanoico, cuya identificación se realizó mediante su tiempo de retención y espectro de masas referenciado en la biblioteca NIST 08, confirmándolo mediante el uso de estándares cromatográficos.

Resultados: La mayoría de los VOC analizados se encuentran a concentraciones muy bajas. Únicamente el nonanal muestra diferencias estadísticamente significativas entre los grupos de estudio, depende exclusivamente del hábito de fumar, y es independiente de la cantidad de tabaco consumido, edad y género.

Conclusiones: El hallazgo de nonanal se asocia al consumo de tabaco, actual o previo. Al ser un producto secundario de la destrucción de la membrana celular, su hallazgo probablemente muestra daño celular en personas fumadoras y permanece una vez cesado el hábito.

© 2013 SEPAR. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Introduction

Tobacco smoking is a risk factor for neoplasms such as lung cancer, chronic obstructive pulmonary disease and cardiovascular diseases, among others.

In most of these diseases, a high degree of oxidative stress occurs primarily due to the amount of free radicals and reactive oxygen and nitrogen species originating in tobacco smoke. Oxidative stress causes irreversible structural and functional cell damage, including changes in essential proteins, lipid peroxidation, DNA strand breakage and nitrogenous base modification, a large increase in intracellular free calcium and in some cases, apoptosis or necrosis.

Of all the changes caused by the increase in oxidative stress, lipid peroxidation generates a large number of metabolites from oxidation reactions catalyzed by enzyme systems associated with cytochrome P450. These include, among others, alkanes, alkenes, aldehydes and carboxylic acids, all of low molecular weight and volatile, known as volatile organic compounds (VOC), which can be excreted via the airways. The presence of these substances in exhaled air indicates oxidative stress in the airways and lungs; for example, linear aldehydes, such as hexanal, heptanal and nonanal compounds are end metabolites of the lipid peroxidation process of fatty acids ω3 and ω6, principal components of the phospholipids that form part of the cell membranes.

The aim of this study was to determine whether smoking influences the production of VOC in exhaled air as a result of oxidative stress in a clinically healthy population. Other compounds producing oxidative stress in biological macromolecules were not studied.

Patients and Methods

Study Design and Definition of Groups

This was a case–control study with consecutive, non-probability sampling. A total of 89 subjects, all volunteers from the staff of the participating centers, including administrative staff, healthcare personnel, physicians, university professors, military personnel and family members, were selected. Possible contamination from their place of residence was not taken into account, as this was not a study objective and, as a variable, would be very complex to define. Three groups were established according to their smoking habits: 35 non-smokers, 24 ex-smokers and 30 active smokers. The status of passive smoker was not considered as a separate group or study parameter.

The inclusion criteria for the 3 groups required personal consent for participation in the study, age over 40 years, and subjects had to be non-smokers, ex-smokers or smokers of over 20 packs a year. Participants were not exposed to special environmental conditions in their work. All had a physical examination and complementary tests, including electrocardiogram, chest X-ray, complete blood count and biochemistry tests and flow/volume loop. None of the participants had signs or symptoms of disease. Exclusion criteria were current or previous lung, heart, endocrine and tumor disease of any system, or refusal to participate in the study. No gender restrictions were applied.

The study protocol was approved by the Clinical Research Ethics Committee of the Hospital Central de la Defensa “Gómez Ulla” (Registry No: 15-404030-5-06-000153).

Compounds Investigated

The compounds studied were linear aldehydes in the order of number of carbons in the molecule as described in the literature consulted (hexanal, heptanal, octanal, nonanal), and carboxylic acids (propanoic acid, hexanoic acid, heptanoic acid, octanoic acid and nonanoic acid), following the same sequence of numbers of carbon atoms as for the aldehydes. Hexanoic, heptanoic and octanoic acids were not detected, so they were not included in the study.

Sampling

After the fasting subject had been rested for one hour without smoking, exhaled air corresponding to the end of the forced vital capacity, the most representative sample of alveolar air possible, was collected. Bio-VOC® Breath Sampler (Markes Int.®) chambers were used and the maneuver was repeated 3 times for preconcentration of the compounds. A sample of ambient air was simultaneously obtained from the room where the subject was studied, to allow comparison of VOC levels in the ambient and exhaled air, with the aim of minimizing possible environmental variations.

After each maneuver, the collected air was transferred to a desorption tube lined with 3 absorbent materials: Tenax TA + Graphitized Carbon Black + Carbonized Molecular Sieve (Markes Int.®). The tube was sealed with a DiffLocks Cap (Markes Int.®) for subsequent analysis. All samples were analyzed within 24 h of collection.

Chromatographic Analysis Method

The analytical technique used was thermal desorption-gas chromatography and mass spectrometry. The analytical method described below was developed by our group.

Tubes containing the samples underwent a desorption process at 250 °C in an Ultra TD® multi-tube autosampler thermal desorption system and a Unity® Thermal desorber with a cold trap (Unity General Purpose Hydrophobic. U-T2GPH [Markes Int.
The compounds were then separated by gas chromatography in a 7890A system, using a DB1-30 m × 0.25 mm × 1 μm chromatographic column with helium as the carrier gas and identified by mass spectrometry with 5975 C VL MSD (Agilent Technologies). The data acquisition mode was in SCAN and the mass interval for detection was set between 40 and 300 m/z (mass/load ratio). The programmed chromatography started at a temperature of 40°C and remained constant for 10 min. It was then increased in increments of 20°C/min until reaching 250°C, where it remained for another 10 min. In the last stage, 300°C was reached at a rate of 30°C/min.

The compounds were identified from their retention time and mass spectra referenced in the NIST 08 mass spectral library, using chromatographic patterns (Fluka®).

The limit of detection (LOD) of a compound was considered to be the minimum concentration at which it could be identified by its mass spectrum, retention time and with a relative abundance of at least 3 times the signal-to-noise ratio, while for the limit of quantification (LOQ), 10 times the signal-to-noise ratio was required. In accordance with both the LOD and LOQ limits, we:

1. **Undetected** compounds: those with a value below the LOD
2. **Unquantified** compounds: compounds detected but which could not be quantified, i.e. the values were between the LOD and the LOQ
3. **Quantified** compounds: those with values above the LOQ

Qualitative data were obtained by integrating the area of the quantifier ion of the mass spectrum for each of the VOC, measured in second counts. Given the possible variability that might exist in the desorption process, the data were normalized according to the area of the 207 ion (corresponding to hexamethylcyclotrisiloxane), used as an internal reference compound, from the desorption tubes.

All quantitative values were calculated using the following formula:

\[
\frac{A_{bc} \text{ VOC}}{A_{bc} \text{ Siloxano}} \times 100
\]

where \(A_{bc}\) is the area under the curve of each compound.

When the ratio of the value in exhaled air/ambient air was greater than 1, the compound was assumed to be endogenous.

The analytical method was validated according to the criteria described in ICH Q2 Validation of Analytical Procedures-2005.7

**Statistical Method**

For descriptive statistics, the mean with standard deviation or median with interquartile range (IQR) and absolute and relative (%) frequencies were used. The following tests were used to determine associations between variables: the Kruskal Wallis test for assessing the association of the different VOC depending on smoking habit; the Mann Whitney test for assessing the association between non-smokers and smokers or ex-smokers and the VOC; and the Spearman linear correlation for assessing nonanal concentration and age.

In all cases, a \(P\)-value <.05 was considered statistically significant and the statistical application used was SPSS® version 15.

**Results**

The characteristics of the study population are listed in Table 1. The sample consists of 89 subjects, all healthy volunteers, with a mean age of 49.3 (9.5) years. With regard to the gender distribution, 47% were men and 53% women. In all, 39% were non-smokers, 34% smokers and 27% ex-smokers, with a tobacco consumption of 30.9

1. **Undetected** compounds: those with a value below the LOD
2. **Unquantified** compounds: compounds detected but which could not be quantified, i.e. the values were between the LOD and the LOQ
3. **Quantified** compounds: those with values above the LOQ

<table>
<thead>
<tr>
<th align="center">Table 1 Characteristics of Study Subjects.</th>
</tr>
</thead>
<tbody>
<tr>
<td align="center"><strong>Age (mean in years), mean (SD)</strong></td>
</tr>
<tr>
<td align="center"><strong>Gender (males/females)</strong></td>
</tr>
<tr>
<td align="center"><strong>Smokers/ex-smokers/non-smokers</strong></td>
</tr>
<tr>
<td align="center"><strong>Smoking habit (packs/year), mean (SD)</strong></td>
</tr>
<tr>
<td align="center"><strong>Ex-smokers</strong></td>
</tr>
</tbody>
</table>

Table 1 Frequency Distribution.

<table>
<thead>
<tr>
<th align="center">Categories</th>
<th align="center">Undetected</th>
<th align="center">Unquantified</th>
<th align="center">Quantified</th>
</tr>
</thead>
<tbody>
<tr>
<td align="center"><strong>Hexanal</strong></td>
<td align="center">64 (72%)</td>
<td align="center">10 (11%)</td>
<td align="center">15 (17%)</td>
</tr>
<tr>
<td align="center"><strong>Heptanal</strong></td>
<td align="center">56 (63%)</td>
<td align="center">5 (6%)</td>
<td align="center">28 (31%)</td>
</tr>
<tr>
<td align="center"><strong>Octanal</strong></td>
<td align="center">72 (81%)</td>
<td align="center">4 (4%)</td>
<td align="center">13 (15%)</td>
</tr>
<tr>
<td align="center"><strong>Nonanal</strong></td>
<td align="center">49 (55%)</td>
<td align="center">9 (10%)</td>
<td align="center">31 (35%)</td>
</tr>
<tr>
<td align="center"><strong>Propanoic acid</strong></td>
<td align="center">41 (46%)</td>
<td align="center">5 (6%)</td>
<td align="center">43 (48%)</td>
</tr>
<tr>
<td align="center"><strong>Nonanoic acid</strong></td>
<td align="center">74 (83%)</td>
<td align="center">1 (1%)</td>
<td align="center">14 (16%)</td>
</tr>
</tbody>
</table>

The values obtained were classified in 3 categories: (a) quantified, (b) unquantified and (c) undetected, according to the criteria defined in the section Patients and methods. The frequency distribution is indicated in Table 2. Standardized values of VOC concentrations deviated from a normal distribution.

**Table 3** shows the median values (IQR) for each of the VOC analyzed for each group. Statistically significant differences were observed between the 3 groups for the nonanal compound only: non-smokers 0.160 (0.160–0.215); ex-smokers 0.160 (0.160–0.5650); and smokers 2.050 (0.160–10.25) \(P<.04\). Fig. 1 shows the \(P\)-values, based on the Mann Whitney test, comparing the groups two by two.

Data scatter in ex-smokers and smokers is very high, so the values of the first group are included in the results of the second. This was the starting point for pooling tobacco users into a single group (smokers and ex-smokers) for comparison with the group of non-smokers. The comparative study of the data obtained between these two groups also shows statistical significance for the nonanal compound (Mann Whitney test, \(P<.019\)). No differences are observed in the other VOC among the groups mentioned.

A qualitative analysis of detection/no detection of VOC was carried out according to the limits of detection and quantification (LOD and LOQ) established and defined above in the section Patients and methods. The frequency distribution in percentage and \(P\)-values calculated with a \(\chi^2\) test are given in Table 4. Once again, the nonanal compound was the only one that showed statistically significant

![Fig. 1. Graph of nonanal median and interquartile range in the three study groups.](image-url)
differences among the groups (non-smokers, 20%; smokers, 53%; ex-smokers, 29%), *P*=.016.

Since the nonanal compound is associated in our sample with tobacco use, we decided to study the relationship between the amount of tobacco smoked (expressed in packs per year) and the amount of nonanal exhaled. After applying Spearman’s linear correlation test, no correlation was observed between the amount of tobacco consumed and the amount of nonanal exhaled (*P*=.823), nor was any relation found between the nonanal compound and age or sex (males: r=0.05; *P*=.753 and females: r=0.081; *P*=.588).

### Discussion

This study evaluates the differences in the VOC concentration found in exhaled air according to tobacco smoking in our study population, consisting exclusively of healthy volunteers.

According to the results, the only compound associated with tobacco use is nonanal. It is known to be of biological origin, a result which is supported by the fact that smoking causes oxidative stress, generating cell destruction and the appearance of VOC. This is an important consideration if nonanal compounds are to be used as a marker for this phenomenon. The finding of nonanal is independent of age, gender and amount of tobacco smoked, but is related with the status of being a smoker or ex-smoker or never having smoked. It is important to observe that no quantitative correlation was observed between tobacco smoked and nonanal exhaled. These findings were not found in the literature consulted, i.e. the appearance of nonanal depended on the status of being a smoker, but was independent of the amount of tobacco smoked. This compels us to think about the persistence of the cell damage produced by tobacco use, in which a high level of oxidative stress is sustained, even if the subject has quit smoking a long time previously. This sustained cell damage may lead to inflammatory disease or tumor, depending on the cytochrome P450 polymorphism (CYP2E1, CYP1A1), which should be the focus of further research.

A review of the literature shows that the majority of studies on VOC in exhaled air are directed toward the search for markers that allow populations with lung cancer to be differentiated from other non-tumor groups. In these studies, the patient groups were distributed according to smoking habit, pooling, on the one hand, active smokers and on the other, ex-smokers and non-smokers, or without making any clear reference to their status as smokers or ex-smokers. Although VOC determinations were investigated in smokers, ex-smokers and non-smokers in the control group in many of the studies, the results always referred to the findings with regard to lung cancer and chronic obstructive pulmonary disease. Few publications however compare VOC in exhaled air in smokers and non-smokers exclusively in a healthy population, allowing the evaluation of the importance of exposure to tobacco smoke as a factor in increased oxidative stress. Chambers et al. opened an avenue with the detection of VOC in blood, but restricted themselves to the study of aromatic hydrocarbons in tobacco smoke.

A wide range of compounds are described in the literature, ranging from alkanes, aromatic hydrocarbons, other hydrocarbons, alcohols, carbonyl compounds and carboxylic acids to many other families of compounds. Studies such as those of Phillips et al. and Poli et al., which showed high sensitivity and specificity using combinations of different hydrocarbons, were later criticized for not being able to explain the metabolic origin of the proposed compounds. A critical analysis of the results of the studies cited above by Hórvath et al. is of interest in this regard.

At present, there is a tendency toward the study of compounds oxygenated as a result of the peroxidation process of the cell components. Various authors, in the most recent papers, already include aldehydes and ketones, among others, as possible markers of oxidative stress. Our study focused on aldehydes and carboxylic acids (as described in the section Patients and methods), due to the high incidence of these 2 groups of compounds in many determinations previously performed. These compounds, which are of known origin, are intermediate and end metabolites of lipid peroxidation processes of the fatty acids that form part of the cell membranes. For example, hexanal, heptanal and nonanal are products of the oxidation of fatty acids ω3 and ω6 (polysaturated fatty acids [PUFA]), principal components of phospholipids. With respect to propanoic acid, this is a natural anti-inflammatory, and nonanoic acid is the result of the complete oxidation of aldehyde in a very oxidizing environment.

It should be remembered that VOC not only have a systemic origin but they may also be exogenous, originating from ambient air or tobacco smoke. Distinguishing compounds of endogenous origin from those of exogenous origin is fundamental for this type of study, in which the collection and handling of the sample before analysis is of utmost importance. Various methodologies have been proposed in this respect. Phillips et al. proposed a simple mathematical calculation, subtracting the ambient concentrations from the normalized concentrations in exhaled air. This approach has been questioned, as it does not explain the complexity of the pulmonary absorption and subsequent exhalation of exogenous compounds after these have been metabolized in the form of different VOC. Another method proposed was to have the patient breathe pure, VOC-free air for a period of between 4 and 30 min. Although this was considered a good method, it has not been followed in most studies.

Several authors, including ourselves, carried out sample collection using a Bio-VOC (as described in the section Patients and

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Median and Interquartile Range of Each VOC in Each Group and P-value.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median and Interquartile range</td>
</tr>
<tr>
<td></td>
<td>Non-smokers=35</td>
</tr>
<tr>
<td>Hexanal</td>
<td>0.095 (0.095–0.200)</td>
</tr>
<tr>
<td>Heptanal</td>
<td>0.055 (0.055–0.930)</td>
</tr>
<tr>
<td>Octanal</td>
<td>0.075 (0.075–0.750)</td>
</tr>
<tr>
<td>Nonanal</td>
<td>0.160 (0.160–0.215)</td>
</tr>
<tr>
<td>Propanoic acid</td>
<td>0.175 (0.105–2.930)</td>
</tr>
<tr>
<td>Nonanoic acid</td>
<td>0.135 (0.135–0.260)</td>
</tr>
</tbody>
</table>

* Kruskal Wallis test.

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Comparison of Percentage According to Smoking Habit, in Which Markers Were Detected in the Different Groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-smokers, %</td>
</tr>
<tr>
<td>Hexanal</td>
<td>11.4</td>
</tr>
<tr>
<td>Heptanal</td>
<td>25.7</td>
</tr>
<tr>
<td>Octanal</td>
<td>8.6</td>
</tr>
<tr>
<td>Nonanal</td>
<td>20.0</td>
</tr>
<tr>
<td>Propanoic acid</td>
<td>51.4</td>
</tr>
<tr>
<td>Nonanoic acid</td>
<td>20.0</td>
</tr>
</tbody>
</table>

* Pearson’s χ² test.
methods), which allows the collection of air from the anatomical dead space, with the aim of avoiding contamination of the sample with non-alveolar air as far as possible. To distinguish exogenous compounds from endogenous compounds, a sample of ambient air was obtained at the same time and both samples were compared. Instead of using the approach of Phillips et al., mentioned above, we standardized the concentration to the internal standard, used as a correction factor for the variability of the tube desorption process. Any compound with an exhaled/ambient ratio greater than 1 was assumed to be endogenous, since this gave us the assurance that we were dealing with a compound of endogenous origin. However, we are aware that with this method, we increase the specificity but reduce the sensitivity.

To conclude, the quantitative or qualitative evaluation of nonanal in exhaled air is associated with tobacco use. The other compounds studied do not show any correlation with smoking. There is a lack of standardization in the methodology for sample collection.

Funding

This study was funded by the Instituto de Salud Carlos III (PI07/1116), Neumomadrid 2008 and SEPAR 2010 (Registry No.: 881). The authors declare that there is no conflict of interest with any of the companies whose products or services may be discussed in this article.

Conflict of Interests

The authors declare that there was no conflict of interest.

Annex

Study Group Members: José Luis Álvarez-Sala Walther – Hospital Clínico San Carlos, Universidad Complutense de Madrid; Rodolfo Álvarez-Sala Walther – Hospital Universitario La Paz, Universidad Autónoma of Madrid; Eva Arias Arias – Hospital 12 of October; Manuel Caamaño Somoza – Universidad Complutense of Madrid; María Vicenta García Rosado – Hospital Central de la Defensa Gómez Ulla; Francisco Roig Vázquez – Hospital Infanta Elena, Madrid; Francisco Villegas-Fernández – Hospital Central de la Defensa Gómez Ulla – Universidad de Alcalá of Henares; Gema Rodríguez Trigo – Hospital Clínico San Carlos, Madrid; José Luis López Colón – Instituto of Toxicología de the Defensa, Madrid.

References

7. ICH Q2 Validation of Analytical Procedures; 2005.