Metabonomic analysis of Exhaled Breath Condensate in adults by NMR spectroscopy

Guglielmo de Laurentiis\textsuperscript{1,2,5}, Debora Paris\textsuperscript{3,5}, Dominique Melck\textsuperscript{3,5}, Mauro Maniscalco\textsuperscript{1}, Serafino Marsico\textsuperscript{2}, Gaetano Corso\textsuperscript{4}, Andrea Motta\textsuperscript{3,*} and Matteo Sofia\textsuperscript{1,*}

\textsuperscript{1}Dept. of Respiratory Medicine, AO Monaldi, University of Naples “Federico II”, Naples (Italy); \textsuperscript{2}Dept. of Respiratory Medicine, AO Monaldi, Seconda Università di Napoli, Naples (Italy); \textsuperscript{3}Institute of Biomolecular Chemistry, Consiglio Nazionale delle Ricerche, Pozzuoli, Naples (Italy); \textsuperscript{4}Dept. of Biomedical Sciences, University of Foggia, Foggia (Italy).

\textsuperscript{5}GdL, DP and DM contributed equally to this work.

Running title: Metabonomic analysis of EBC by NMR

Address correspondence to:
Guglielmo de Laurentiis, M.D.,
Dept. of Respiratory Medicine, AO Monaldi,
Via L. Bianchi, 80131 Naples, Italy
Phone: +39 0817062377;
Fax: +39 0815453213;
e-mail: g.delaurentiis@hotmail.it
Abstract

Background: Exhaled Breath Condensate (EBC) is a non-invasive method to study airway lining fluid. Nuclear Magnetic Resonance (NMR) spectroscopy can provide biochemical profiles of metabolites in biological samples.

Objectives: To validate the NMR-metabonomic analysis of EBC in adults, assessing the role of preanalytical variables (saliva and disinfectant contamination), and its potential clinical feasibility.

Material and Methods: 36 paired EBC and saliva samples, obtained from healthy subjects (HS), laryngectomized and chronic obstructive pulmonary disease (COPD) patients, were analyzed by means of $^1$H-NMR spectroscopy followed by principal component analysis (PCA). The effect on EBC of disinfectant, used for reusable parts of the condenser, was assessed after different washing procedures. To evaluate intra-day-repeatability, 8 subjects were asked to collect EBC and saliva twice within the same day.

Results: All NMR saliva spectra were significantly different from corresponding EBC samples, EBC taken from condenser washed with recommended procedures invariably showed spectra perturbed by disinfectant. Each EBC sample clustered with corresponding samples of the same group, while presenting intergroup qualitative and quantitative signal differences (94%).

Conclusions: The NMR-metabonomic approach could identify the metabolic fingerprint of EBC in different clinical sets of data. Moreover, metabonomics of EBC in adults can discriminate potential perturbations induced by preanalytical variables.

Keywords: Exhaled Breath Condensate, Metabonomics, Nuclear Magnetic Resonance, Principal Component Analysis, Saliva, Chronic Obstructive Pulmonary Disease.

Final Word Count: 3725
Introduction
Exhaled breath condensate (EBC) is a simple, non invasive and useful tool to study the biochemical and inflammatory molecules in the airway lining fluid (1). Obtained by cooling exhaled air under condition of spontaneous breathing, EBC contains predominantly water vapor and collects volatile and non-volatile substances from the lower airways (2). As such, it can also be considered a matrix for analysis of environmental toxicant, and for evaluation of exposure monitoring (3). Very few data are available on EBC metabolite composition, often analyzing single inflammatory molecules by ELISA and spectroscopic methods.

Nuclear Magnetic Resonance (NMR) studies molecules by recording the interaction of radiofrequency electromagnetic radiation with the nuclei placed in a strong magnetic field. A single nucleus in a molecule can be “observed” by monitoring the corresponding line in an NMR spectrum, and the various parameters of that line (frequency, splitting, linewidth and amplitude) can be used to determine the molecular structure, conformation and dynamics of biomolecules. In principle, assignment (i.e., identification) of NMR lines could be possible by comparing the observed chemical shifts (i.e., the positions in a spectrum) with reference data, spreading out the spectra peaks in two dimensions. Recently, NMR has been applied to biofluids to probe the metabolic status (4-7), and to investigate different diseases (8-11).

The presence of a discriminating element in a series of spectra (i.e., the spectroscopic space) is often undetectable due to the inherent complexity, and is better highlighted by multivariate analysis (Principal Component Analysis, PCA) that carefully identifies hidden phenomena and trends in ensembles of spectra (12). The application of PCA to a group of spectra can immediately rationalize if all spectra behave similarly or fall apart into different groups.

Recently, EBC of asthmatic children has been investigated by NMR and statistical analysis (13). To date there are several recommendations on the methodological approach to EBC collection, but its standardization is not completely defined, as most inflammatory mediators, obtained through tracheostomies, are similar to those collected in the mouth (14, 15).

The aims of our study was a) to validate the NMR-metabonomic approach to analysis of EBC in adults, assessing the role of preanalytical variables (saliva and disinfectant contamination) potentially influencing EBC and evaluating the stability and reproducibility of samples; b) to evaluate the possibility to discriminate healthy subjects (HS) from patients with airway disease.
Materials and methods

Subjects

A total of 36 paired EBC samples were collected from the following groups of subjects: 12 HS (9 males, mean age 55.6 ± 7.2 years), 12 laryngectomized patients (9 males, mean age 60.2 ± 6.2 years), and 12 patients affected by chronic obstructive pulmonary disease (COPD, 11 males, mean age 64.9 ± 5.7 years).

All HS subjects were non smokers, while the laryngectomized patients (who provide sample through a stoma, bypassing the pharynx entirely) and the COPD patients were ex smokers (smoking free time from at least 24 months). All subjects presented no occupational or other pronounced exposure to organic solvents. The laryngectomized patients had been previously treated by laryngectomy for laryngeal carcinoma > 1 year before (range 12-18 months) and they did not refer history of chronic respiratory disease or recurrent exacerbations. COPD patients have received diagnosis in the past according to Global Obstructive Lung Disease (GOLD) guidelines (16). The COPD anthropometric characteristics are summarized in Table 1. None was on regular systemic or inhaled corticosteroid treatment. They were asked not to use long acting beta-2 agonist and anticholinergic agents for at least 12 and 24 hours respectively before EBC collection.

All subjects were free from upper and/or lower airway infection from at least 4 weeks before the EBC collection. They refrain from food intake at least 4 hours from the test and drinking alcoholics for at least 18 hours before EBC collection. In laryngectomized patients lower respiratory tract secretions were actively managed by self-suctioning and cleaning before each EBC collection.

To assess within-day repeatability, 8 subjects (4 HS and 4 COPD patients) were asked to collect EBC and saliva twice within the same day (time 0 and 12 hours after).

All subjects gave the informed consent and the study protocol was approved by the Ethics Committee of the Monaldi Hospital.

EBC sampling

EBC was collected using an EcoScreen condenser (Jaeger, Wyrzburg, Germany) as previously reported (17). Briefly, all subjects breathed through a mouthpiece (laryngectomized patients provided samples through the stoma) and a two-ways nonrebreathing valve, which also served as a saliva trap, at normal frequency and tidal volume, while sitting comfortably and wearing a nose-clip, for a period of 15 minutes. They maintained a dry mouth during collection by periodically swallowing excess saliva.

Condensate samples (3-4 ml) were immediately transferred into glass vials of 10 ml volume, closed with 20-mm butyl rubber lined with PTFE septa, and crimped with perforated aluminum seals.
Volatile substances possibly deriving from extrapulmonary sources (18, 19) were removed by a gentle stream of nitrogen before sealing. Nitrogen was applied for variable time (1, 3, 5, 10, 15 and 20 minutes); no difference was observed with spectra acquired after 1-minute nitrogen exposure, but since such interval appeared to be too short to avoid systematic errors, we chose a 3-minute interval. Nitrogen was used because concentration of volatile solutes in EBC is dependent on their distribution between the saliva, exhaled air and droplets, and the condensate, which can be altered by multiple factors including minute ventilation, salivary pH, solubility, temperature, and sample preparation (20). Therefore, spectral differences may depend upon uncontrollable variables that prevent reliable quantification. The nitrogen stream also removes oxygen from solutions. Such a procedure, used for NMR protein structure determination (21), together with freezing of sealed samples in liquid nitrogen, immediately “quenches” metabolism at the collection time, and prevents any metabolic decay (8). Samples were then stored at -80°C until NMR analysis. We avoided drying of the samples to circumvent irreversible solute precipitation, and/or formation of insoluble aggregates, which we observed upon dissolving the dried condensate for NMR measurements.

Preanalytical preparation of EBC condenser reusable parts
Before and after collection of each EBC sample, the reusable parts (valve, salivary trap and lamellar condenser) were disinfected for 15 minutes using a solution of a 1.5% freshly prepared chemical agent (Descogen™, FILT GmbH, Berlin, Germany), and repeatedly flushed with water following the manufacturer’s guidelines. To completely eliminate the disinfectant, parts already disinfected and washed were thoroughly rinsed for 15 minutes with pure grade ethanol (96%), thereafter exhaustively soaked with deionized distilled water for 15 minutes and dried under vacuum at 50°C.

Salivary collection
In the same day of EBC collection, we also took a salivary sample. To avoid any interference from exogenous agents into the oral environment, we asked the patients to collect all saliva available (ca. 2-4 ml), i.e., (“whole”) saliva expectorated from the mouth, into a plastic universal tube immediately after waking in the morning. According to Silwood et al. (11), each patient was requested to refrain completely from oral activities (i.e., eating, drinking, tooth brushing, oral rinsing, smoking, etc.) during the short period between awakening and sample collection (< 5 minutes). Each collection tube contained 15 µmol sodium fluoride, sufficient to ensure that metabolites are not generated or consumed via the actions of bacteria or bacterial enzymes present in whole saliva during periods of sample preparation and/or storage (11). Specimens were
transported to the laboratory on ice and immediately centrifuged (12000 rpm at 4°C for 15 minutes) on their arrival to remove cells and debris, and the a gentle nitrogen gas flow was applied for ca. 5 minutes to supernatants, which were then stored at -80°C until measurements (11).

The 1H NMR profiles of salivary supernatant specimens subjected to analysis immediately after collection into the fluoride-containing tubes and rapid centrifugation were compared with those of the same samples stored as described above, and no differences were discernible, i.e., none of the criteria investigated changed significantly during these periods of storage.

**Sample preparation for NMR analysis**

EBC samples were rapidly defrosted. To provide a field frequency lock, 70 µl of a D2O solution [containing 1 mM sodium 3-trimethylsilyl [2,2,3,3-2H4] propionate (TSP) as a chemical shift reference for 1H spectra, and sodium azide at 3mM], were added to 630µl of condensate reaching 700 µl of total volume. Saliva samples were rapidly defrosted and 70µl of reference standard solution (D2O-TSP) were added to 630µl of sample.

**NMR measurements**

One-dimensional (1D) spectra were recorded on a Bruker Avance spectrometer operating at a frequency of 600.13 MHz (1H) and equipped with a TCI CryoProbe™, at a probe temperature of 27°C. The water resonance was suppressed by using the noesypresat pulse sequence, called noesypr1d in the Bruker language. It has the form –RD-90°-t-90°-tm-ACQ, where RD is a relaxation delay, t a short delay, 90° represents a 90° RF pulse, tm the mixing time, and ACQ the data acquisition period. In our acquisition conditions, the carrier frequency (O1) value was set on the water resonance, the saturation power to 62 dB, t = 4 µs, tm =100 ms, the spectral amplitude (SW) to 7002.8 Hz, TD = 16K, RD = 2.0 s, and number of transients to 256. This results in a total acquisition time of 14 minutes per sample. For processing, a line broadening of 0.6 Hz was applied and an SI of 32K was used. Spectra were referred to TSP assumed to resonate at δ = 0.00 ppm.

**Statistical analysis**

High-resolution 1H-NMR spectra were automatically data reduced to 200 integral segments (“buckets”), each of 0.02 ppm, using the AMIX software package (Bruker Biospin, Germany). The resulting integrated regions were imported into the SIMCA package (Umetrics, Umea, Sweden) and used for statistical analysis and pattern recognition. Before pattern recognition analysis each integral region is usually normalized to the sum of all integral region of each spectrum; however, because of the presence of contaminant peaks, each bucket was normalized to the TSP peak of
known concentration for a reference region comprised between 0.014 and -0.014 ppm. The correctness of the approach was tested by comparing the results with those obtained by referring to the sum of all integral region of each contaminant-free spectrum. No significant difference was observed between the two approaches, therefore granting reliable pattern recognition analysis with normalization to TSP. Data were preprocessed with the Centering scaling and then processed with PCA and Partial Least Squares-Discriminant Analysis (PLS-DA).
Results

Spectral differences between EBC and saliva

Figure 1 represents spectra of saliva (traces A, B and C) and EBC samples (traces D, E and F) from an HS (A and D, respectively), a laryngectomized patient (B and E) and a COPD patient (C and F). Saliva spectra were highly different from corresponding EBC samples, and were notably dissimilar among them: a visual examination establishes a correspondence between spectra A (HS) and B (laryngectomized), but a difference with the COPD spectrum, which shows sharper lines. The most intense signals in the 0.0-3.2 ppm region of saliva were assigned (see caption to Figure 1) by resorting to literature data (11, 22). We identified resonances from leucine, propionate, lactate, threonine, alanine, acetate, glutamate, glutamine, pyruvate, succinate, lysine, choline, phosphorylcholine, and taurine. Signals between 3.3 and 6.0 ppm originate from carbohydrates and are virtually absent in the EBC spectra.

Compared to saliva, EBC spectra present fewer signals, and, as observed for saliva, COPD trace F appears to be different from HS (trace D) and laryngectomized (trace E).

Spectral differences between saliva and EBC were verified by resorting to PLS-DA analysis. Because of the complete absence of the carbohydrates signals in the EBC spectrum, the region 5.0 to 3.5 ppm was cut out from all spectra, partitioning the region comprised between 3.5 and 0.8 ppm. Figure 2 shows the score plots of saliva and EBC samples from all subjects. Considering two PLS-DA components, it is possible to obtain a samples’ classification of ca. 95% (samples correctly classified into different regions). In particular, while EBC are all clustered, the saliva samples of HS, laryngectomized and COPD patients are positioned differently from EBC and from each other. Such a separation comes mostly from signals resonating in the within 3.5-2.9 and 2.1-1.7 ppm regions.

EBC and saliva samples collected from 8 subjects twice within the same day (time 0 and after 12 hours) demonstrated good within-day repeatability, showing the spectra no evident difference of resonances.

Effects of disinfectant contamination on EBC spectra

Figure 3 shows the 1D ¹H-NMR spectrum of Descogen™ (trace A) with the representative spectra of EBC samples contaminated by the disinfectant because of insufficient washing time (spectra B and C). To completely eliminate the disinfectant, parts already disinfected and washed were thoroughly rinsed for 15 minutes with pure grade ethanol (96%), thereafter exhaustively soaked with deionized distilled water for 15 minutes and dried under vacuum at 50°C.
The resonances of the “saline” components of the disinfectant (citric acid, at 2.66 ppm in the Descogen™ spectrum A, and pentapotassium bis(peroxymonosulfate) bis(sulfate), highly soluble in water) disappeared completely after partial washing (15 minutes, spectrum B). However, minor unknown components, as those originating the signals in the 8.2-7.3 and in 1.3-0.7 ppm regions, and the one located at 3.2 ppm, appear to be more persistent even after intense water rinsing (30 minutes, spectrum C). They are completely removed only after the washing procedure using ethanol (spectrum D).

As the perturbation induced by the disinfectant contamination of EBC samples showed visible signals, we examined two different contaminated sets of 12 EBC samples from all COPD patients after partial washing (15 minutes, “high Descogen™”, spectrum B); and after intense water rinsing (30 minutes, “low Descogen™”, spectrum C). Since the region 8.5-7.0 ppm is absent in the “cleaned” EBC spectrum (Figure 3D), as suggested by Carraro and coworkers (4) we used the region 4.5 to 0.5 ppm, and excluded the lactate signals (corresponding to those marked by an asterisk in Figure 3D). Considering two PLS-DA components, we obtained a classification that was around 72%, with high-Descogen™ (empty boxes) and low-Descogen™ (filled boxes) EBC samples classified in two wide regions (Figure 4A). This suggests that the presence of the disinfectant at variable concentration affects the interpretation and the statistical analysis of the samples. However, if we “ignore” the presence of contaminant by a careful selection of the spectral regions to be used for statistical analysis, it is possible to correctly classify the samples. In fact, selecting only the Descogen™-free regions of the spectra (3.5-2.9 and 2.1-1.7 ppm) all the samples are correctly classified (94%) in a single region (Figure 4B), as it would have been expected from “clean” (i.e., totally Descogen™-free) EBC samples.

**EBC spectral discrimination between healthy subjects, laryngectomized and COPD patients**

The 3.5-1.7 ppm region of “clean” (i.e. Descogen™-free) EBC samples was used to investigate the metabolites characterizing EBC. Figure 5 depicts representative spectra of a healthy (trace A), laryngectomized (trace B) and COPD (trace C) subjects. Although the region contains few signals, they specifically characterize each patient subset, showing both quantitative (signals intensity) and qualitative (signals absence/presence) differences. Differences in intensity are shown by the acetate, propionate, pyruvate (present in COPD spectrum C and very intense in healthy spectrum A, but barely visible in laryngectomized spectrum B); succinate (small in healthy spectrum A, bigger in laryngectomized spectrum B but absent in COPD spectrum C); glutamine (only present in healthy spectrum A), the singlet at 3.03 ppm (only present in COPD spectrum C), most likely originating from N-CH₃ of creatine/creatinine; choline and phosphorylcoline, absent in COPD...
spectrum C; methanol; and trimethylamine-N-oxide (TMAO) present in healthy spectrum A, barely seen in laryngectomized spectrum B, and absent in COPD spectrum C. All these differences prompted a clear discrimination of HS, laryngectomized and COPD patients in three separate groups (Figure 6).
Discussion
This study demonstrates for the first time that NMR-based metabonomics can be used to analyze EBC samples from adults, allowing a clearcut separation between healthy subjects and patients with airway disease. Although less sensitive than ELISA and mass spectrometry, NMR requires minimal sample preparation with a rapid acquisition time (ca. 10-15 minutes), and has a high degree of sensitivity (≤ µmol/l). Furthermore, it is nondestructive and allows a complete detection of observable metabolites (“sample metabolic fingerprint”) at a reasonable cost.

Our data show that saliva is significantly different from the EBC samples and that the presence of identical metabolites in EBC and in saliva does not hamper discrimination. By selecting the 3.5-0.8 ppm region (therefore excluding the carbohydrate signals absent in EBC), saliva spectra clearly differ from EBC (Figure 2), notwithstanding the presence of some common metabolites (leucine, lactate, propionate, acetate, etc.).

EBC standardizing guidelines (2) indicate that it is reasonable to assume that there is some degree of oral contamination of EBC, as saliva contains many of the mediators that are also present in the lower airways. Contamination of EBC is often proved by measuring the amylase level; such a test is not specific and a negative signal does not completely exclude minute contribution from the mouth. To date, there are no data comparing the metabolic saliva composition and a lower airway derivate such as EBC, mainly because condensate samples have been screened for single, specific biomarkers, and not as a whole. Indeed, combined saliva and EBC analysis by metabonomics method has been recently advocated (19). In light of these assumptions, we also examined EBC from laryngectomized patients, which may represent a true saliva-free material from the lower airways, showing that in those subjects all saliva spectra strictly differed from corresponding EBC samples. Importantly, all EBC and saliva collected twice within the same day (12 hours after) showed good repeatability within-day (Figure 6). Taken together, the above data suggest that saliva contamination may play a minor role in the interpretation of EBC by NMR-based metabonomics.

We also considered the influence of external contaminant, as the International Consensus on EBC recommends special care in the disinfection of reusable parts of the condensers (1). Upon standard cleaning, all EBC spectra presented signals corresponding to unknown inactive substances of the disinfectant. They persisted even after strong and repeated water soaking, and the presence of variable disinfectant concentration upon different cleaning levels may render classification less effective. We observed complete removal of the disinfectant signals after washing the reusable parts with ethanol 96 %, and then rinsing thoroughly with distilled water for 15 minutes. So, we spiked EBC samples by partially washing the apparatus with water, after treatment with freshly
prepared Descogen™, obtaining different degree of EBC contamination. Since the citric acid signals in spectrum 3B were absent, it is important to underline that the potentially toxic “saline” components of the disinfectant are easily removed from the condenser apparatus after water washing. However, the persistence of interfering residual external contaminants is crucial for a correct EBC analysis. There are no data on the influence of residual disinfectants agents of reusable parts of EBC condensers. We did not evaluate the influence of residual Descogen™ on reported biomarkers levels by ELISA method, but we suggest that the potential role of external contamination on the variability of some biomarkers (23, 24) should be evaluated.

Significantly, by selecting specific regions of EBC spectra for statistical analysis, we obtained an efficient discrimination of samples. Although separation between HS and COPD can be achieved by either FEV₁ measurements or clinically, we evaluated the capability of NMR-based metabonomics to separate EBC subjects with airway diseases (COPD) from subjects without respiratory diseases. Five NMR signals appear to differentiate “respiratory” (COPD) from “non respiratory” (HS and laryngectomized) subjects. As a comparison, Carraro et al. (4) reported the single acetate signal variation as distinctive in asthmatics children with respect to controls. They hypothesized that acetate increase might be related to increased acetylation of proinflammatory proteins in the extracellular space in the airway environment. Furthermore, they found that peaks in 3.2-3.4 ppm regions of the NMR spectrum of asthmatic children were probably related to oxidized compound. Heili-Frades et al. in an abstract form (25) have reported preliminary data on significant variations between NMR EBC spectra of normal and pathological cases with implications for correlative studies using spectral and clinical classification.

In the present study, by comparing EBC from “respiratory” (COPD) patients and “non respiratory” (HS and laryngectomized) subjects, together with acetate, we found four additive signal variations, which likely included the methoxy compounds. Such variations could speculatively derive from an increased oxidative stress that represents a hallmark of COPD, and are usually investigated on EBC by measuring a limited number of markers (23-24). Finally, the comparison between HS, laryngectomized and COPD EBC samples showed a clearcut difference (Figure 5) of COPD from the other subjects. Figure 6 depicts a significant statistical difference along t[1] of COPD with HS and laryngectomized patients, who are less separated along t[2]. This could be interpreted by the fact that laryngectomized patients before or after surgery were not labeled as COPD; furthermore, mild airflow limitation was detected only in a few subjects (data not shown).

We are aware that our study relies on the limited number of subjects without a prospective test sample of different groups of subjects, followed by an external validation in a second set of subjects. However, we are collecting more NMR data on EBC so as to build a data bank to be used,
together with other techniques, for prognostic and/or diagnostic purpose. Moreover, the observed metabolites were not identified at this stage, and we are currently characterizing all resonances by using homo- and heteronuclear twodimensional spectra, and the results will be reported in due course.

In conclusion, NMR-based metabonomics can safely be applied to EBC in adults, allowing an unambiguous definition irrespective of natural and/or artificial contaminants. In particular, we have reported that NMR spectra of EBC, collected with a device using a salivary trap, do not show the presence of saliva signals. Furthermore, for the disinfectant medium currently used, a careful selection of the NMR region allows a clear statistical classification of samples, even for contaminated EBC samples. Finally, the reported results suggest that condensate can be efficiently studied as a whole, and that NMR may become a leading diagnostic technique in this field.
Acknowledgments

GdL thanks the Second University of Naples for a fellowship entitled “Fisiopatologia medico-chirurgica del sistema cardio-respiratorio e biotecnologie associate”, during his PhD. DP thanks the Telethon Institute of Genetics and Medicine (Naples, Italy) for a fellowship. We finally thank Emilio P. Castelluccio (IBC-CNR, Pozzuoli, Italy) for computer assistance and Dr. Mauro Mormile (University Federico II, Naples, Italy) for technical support with stomospirometry in laryngectomized patients.
References


Table 1. Anthropometric characteristics of 12 patients affected by Chronic Obstructive Pulmonary Disease (COPD). All results are represented as mean ± SD. FEV$_1$= Forced expiratory volume at first second; BMI= Body mass index; GOLD= Global Obstructive Lung Disease stage. FEV$_1$ data are expressed as post-bronchodilatating inhalation test.

<table>
<thead>
<tr>
<th>n</th>
<th>Age (years)</th>
<th>Sex</th>
<th>BMI (Kg/m2)</th>
<th>FEV$_1$ (%th)</th>
<th>FEV$_1$ (L/sec)</th>
<th>FVC (L/sec)</th>
<th>FEV$_1$/FVC %</th>
<th>GOLD stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64</td>
<td>M</td>
<td>26,6</td>
<td>40</td>
<td>0,95</td>
<td>2,35</td>
<td>40,3</td>
<td>III</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>M</td>
<td>24,8</td>
<td>70</td>
<td>1,86</td>
<td>3,12</td>
<td>59,5</td>
<td>I</td>
</tr>
<tr>
<td>3</td>
<td>66</td>
<td>M</td>
<td>27,1</td>
<td>50</td>
<td>1,35</td>
<td>2,98</td>
<td>45,1</td>
<td>III</td>
</tr>
<tr>
<td>4</td>
<td>69</td>
<td>M</td>
<td>29,7</td>
<td>61</td>
<td>1,59</td>
<td>3,15</td>
<td>50,3</td>
<td>II</td>
</tr>
<tr>
<td>5</td>
<td>75</td>
<td>M</td>
<td>33,5</td>
<td>30</td>
<td>0,90</td>
<td>2,15</td>
<td>41,9</td>
<td>IV</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>M</td>
<td>32,4</td>
<td>40</td>
<td>1,01</td>
<td>2,21</td>
<td>45,8</td>
<td>III</td>
</tr>
<tr>
<td>7</td>
<td>64</td>
<td>M</td>
<td>25,0</td>
<td>50</td>
<td>1,27</td>
<td>3,15</td>
<td>40,2</td>
<td>III</td>
</tr>
<tr>
<td>8</td>
<td>66</td>
<td>M</td>
<td>25,9</td>
<td>70</td>
<td>1,74</td>
<td>2,89</td>
<td>60,1</td>
<td>I</td>
</tr>
<tr>
<td>9</td>
<td>71</td>
<td>M</td>
<td>26,0</td>
<td>38</td>
<td>0,90</td>
<td>1,89</td>
<td>47,4</td>
<td>III</td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>F</td>
<td>30,1</td>
<td>34</td>
<td>0,90</td>
<td>2,16</td>
<td>41,7</td>
<td>III</td>
</tr>
<tr>
<td>11</td>
<td>56</td>
<td>M</td>
<td>24,5</td>
<td>30</td>
<td>0,89</td>
<td>2,35</td>
<td>38,0</td>
<td>IV</td>
</tr>
<tr>
<td>12</td>
<td>58</td>
<td>M</td>
<td>25,3</td>
<td>60</td>
<td>1,75</td>
<td>3,21</td>
<td>54,6</td>
<td>II</td>
</tr>
</tbody>
</table>
Figures Legends

Figure 1. Representative one-dimensional $^1$H-NMR spectra of saliva and EBC samples from healthy (A and D, respectively), laryngectomized (B and E), and COPD (C and F) patients. The group of signals centered at 3.8 ppm in saliva spectra originates from carbohydrates and are not visible in the corresponding EBC spectra. Resonances assignment is as follows: leucine $\delta$CH$_3$s (triplet) at 0.96 ppm; propionate $\beta$CH$_3$ at 1.04 ppm (triplet) and $\alpha$CH$_2$ at 2.19 ppm (quartet); lactate $\beta$CH$_3$ at 1.32 ppm (doublet) and $\alpha$CH at 4.11 ppm (quartet); the threonine $\gamma$CH$_3$ at 1.36 ppm (doublet); alanine $\beta$CH$_3$ at 1.47 ppm (doublet) and $\alpha$CH at 4.20 ppm (quartet); and acetate CH$_3$ (singlet) at 1.93 ppm; $\beta$CH$_2$ of glutamate and glutamine at 2.10 ppm (multiplet); $\beta$CH$_3$ of pyruvate at 2.37 ppm (singlet); $\alpha,\beta$CH$_2$ of succinate at 2.41 ppm (singlet); $\varepsilon$CH$_2$ of lysine at 3.06 (triplet); N-CH$_3$s of choline at 3.16 ppm and of phosphorylcholine at 3.23 ppm (both singlets); N-CH$_3$ of taurine at 3.23 ppm (triplet).
Figure 2. Partial Least Squares-Discriminant Analysis (PLS-DA) scores discrimination for EBC (empty triangles, laryngectomized; filled triangles, HS; and filled circles, COPD) and saliva (empty squares, laryngectomized; filled squares, HS; empty circles, COPD). All variables have been used and two PLS components are retained in the model, obtaining a classification of ca. 95%. The region 5.0 to 3.5 ppm, containing the carbohydrates signals, was cut out from the bucketing, and only the signals comprised between 3.5 and 0.8 ppm were analyzed.
Figure 3. Contamination of EBC samples by Descogen™. (A) $^1$H-NMR spectrum of Descogen™, compared with spectra of EBC samples after partial washing (15 minutes, B), and intense water rinsing (30 minutes, C). Contamination is completely removed after the washing procedure using ethanol (D). The acetate signal at 1.93 ppm is cut in all EBC spectra, and an asterisk marks the lactate resonances. The vertical scale of the Descogen™ spectrum is $\frac{1}{4}$ of the other spectra.
Figure 4. PLS-DA scores discrimination for contaminated EBC samples at different washing time (A), and selecting the 3.5 to 2.9 and 2.1 to 1.7 ppm regions for statistics (B). In both plots filled squares represent “high Descogen™” (15-minute rinsing, as from spectrum B in Figure 3), while empty squares refer “low Descogen™” (30-minute rinsing, as from spectrum C in Figure 3). Exclusion of the most intense Descogen™ signals in (B), allows a classification of ca. 94%, as it would have been expected from “clean” EBC samples.
Figure 5. Representative $^1$H-NMR spectra of contaminant-free EBC samples from healthy (A), laryngectomized (B), and COPD (C) patients. The acetate singlet at 1.93 ppm is cut by a horizontal bar. Differences in the signals are described in the text. Signals identification is as follows: acetate $\beta$CH$_3$ (singlet) at 1.93 ppm; propionate $\alpha$CH$_2$ at 2.19 ppm (quartet); pyruvate $\beta$CH$_3$ (singlet) at 2.37 ppm; succinate $\alpha,\beta$CH$_2$ (singlet) at 2.41 ppm; glutamine $\gamma$CH$_2$ (multiplet) at 2.45 ppm; singlet at 3.03 ppm most likely originating from N-CH$_3$ of creatine/creatinine; choline and phosphorylcoline N-CH$_3$s (singlets) respectively at 3.16 and 3.23 ppm; methanol CH$_3$ at 3.37 ppm (singlet); and trimethylamine-N-oxide (TMAO) N-CH$_3$ (singlet) at 3.44.
Figure 6. PLS-DA scores discrimination for contaminant-free EBC samples. Filled squares represent healthy subjects; empty squares refer to laryngectomized patients and empty circles to COPD patients. Two PLS-DA components afforded a clear classification (ca. 94%), with all
samples correctly classified into three regions. Vertical and horizontal bars for empty circles and filled squares refer to samples collected in duplicate.