Anal Bioanal Chem (2010) 398:3043–3050
DOI 10.1007/s00216-010-4238-y

ORIGINAL PAPER

Chemical characterization of exhaled breath to differentiate between patients with malignant plueral mesothelioma from subjects with similar professional asbestos exposure

G. de Gennaro · S. Dragonieri · F. Longobardi · M. Musti · G. Stallone · L. Trizio · M. Tutino

Received: 12 July 2010 / Revised: 18 September 2010 / Accepted: 21 September 2010 / Published online: 6 October 2010

Abstract Malignant pleural mesothelioma (MPM) is an aggressive tumour whose main aetiology is the long-term exposure to asbestos fibres. The diagnostic procedure of MPM is difficult and often requires invasive approaches; therefore, it is clinically important to find accurate markers for MPM by new noninvasive methods that may facilitate the diagnostic process and identify patients at an earlier stage. In the present study, the exhaled breath of 13 patients with histology-established diagnosis of MPM, 13 subjects with long-term certified professional exposure to asbestos (EXP) and 13 healthy subjects without exposure to asbestos (healthy controls, HC) were analysed. An analytical procedure to determine volatile organic compounds by sampling of air on a bed of solid sorbent and thermal desorption GC-MS analysis was developed in order to identify the compounds capable of discriminating among the three groups. The application of univariate (ANOVA) and multivariate statistical treatments (PCA, DFA and CP-ANN) showed that cyclopentane and cyclohexane were the dominant variables able to discriminate among the three groups. In particular, it was found that cyclohexane is the only compound able to differentiate the MPM group from the other two; therefore, it can be a possible marker of MPM. Cyclopentane is the dominant compound in the discrimination between EXP and the other groups (MPM and HC); then, it can be considered a good indicator for long-term asbestos exposure. This result suggests the need to perform frequent and thorough investigations on people exposed to asbestos in order to constantly monitor their state of health or possibly to study the evolution of disease over time.

Keywords Biomarkers · Malignant pleural mesothelioma · Exhaled breath · Volatile organic compounds

G. de Gennaro · F. Longobardi · G. Stallone · L. Trizio · M. Tutino (✉)
Department of Chemistry, University of Bari Aldo Moro, 70126 Bari, Italy
e-mail: tutino@chimica.uniba.it

S. Dragonieri
Department of Pulmonology, University of Bari Aldo Moro, 70121 Bari, Italy

M. Musti
Department of Occupational Medicine, University of Bari Aldo Moro, 70121 Bari, Italy

Introduction

Malignant pleural mesothelioma (MPM) is an aggressive neoplasm of mesothelial cell origin that arises mainly from the pleura [1, 2]. The major histological subtypes are epithelial, sarcomatoid and mixed. In addition, osteosarcomatous degeneration within MPM is considered a rare subtype. The majority of MPM cases are associated with asbestos exposure. Occupational asbestos exposure was clearly assessed as the main factor involved in MPM pathogenesis, although asbestos exposure can be environmental [3]. Despite the ban of asbestos usage in developed countries, due to the long latency up to 40 years after asbestos exposure, MPM is nowadays a relevant public health issue since its increasing incidence worldwide over the next 20 years [4]. Moreover, because of the lack of regulations for the use of asbestos in developing countries, MPM will remain a major health problem worldwide for many years.

The diagnostic procedure of MPM is difficult, very often requiring invasive approaches such as thoracoscopy. Therefore, patients are often diagnosed late in the evolution of

the disease when curative treatment is no longer an option [5, 6]. Thus, it is clinically important to find accurate biomarkers for MPM by new noninvasive methods that may facilitate the diagnostic process and identify patients requiring further invasive procedures at an earlier stage [7]. However, the search for biomarkers is hampered by the complex biological characteristics of mesothelioma and the clinical need to distinguish MPM from more frequent pathologies such as benign pleural diseases and metastatic malignancies. Recently, promising diagnostic biomarkers have been proposed in serum and pleural fluid such as osteopontin, soluble mesothelin and megakaryocyte potentiating factor showing a link with the presence of the tumour in patients with MPM [3, 7]. These findings suggest that there are multiple candidate biomarkers that might be captured by diagnostic pattern recognition. Noninvasive analytic methods based on the presence of hundreds of volatile organic compounds (VOCs) in exhaled air could further expand the use of diagnostics. Exhaled air is easily obtained from patients which facilitates repeated sampling of not only the same patient but also larger populations of patients at a lower cost. Human breath contains numerous volatile substances derived both from endogenous metabolism and exposure to ambient vapours and gases and their metabolites. Approximately 3,000 different compounds were detected in human breath; some are correlated with various common disorders like diabetes, heart disease and possibly lung cancer [8–13]. Generally, the composition of different constituents in expired air is representative for blood concentrations resulting from gas exchange at the blood/breath interface in the lungs [14]. Thus, the presence and concentrations of specific volatile organic compounds in expired air are directly linked to their presence in the blood, which is in contact with diseased tissues and organs. Furthermore, metabolites derived from local bacterial infections in the airways can also be detected directly in breath. Pulmonary infections carry a significant risk for people with weak immune systems, especially for long-term and postoperative patients.

Different techniques (online or off-line sampling and analysis) are used for breath analysis [13, 15, 16]. The number of compounds detected in exhaled air and their concentrations vary according to the sampling procedure and the analytical method used. The major VOCs found in the breath of healthy persons (with their typical concentrations in parts per billion by volume, ppbv) are isoprene (10–600), acetone (1–2,000), ethanol (10–1,000) and methanol (150–200 ppbv). All are products of the standard metabolic processes [17]. Selected ion flow tube mass spectrometry and proton transfer reaction time-of-flight mass spectrometry are the most used online techniques for measuring volatile organic compounds in breath samples. The suitability of off-line sampling necessitates the use of appropriate receptacles in which to collect, transport and store gas samples before analyses. In the present work, Tedlar bags were used to collect human breath. The components of the exhaled breath sample were collected using a sorbent-trap containing a bed of a solid sorbent followed by thermal desorption and analysis by gas chromatograph coupled with a mass spectrometer (TD-GC-MS). The aim of the study was to discriminate among exhaled breath of patients with MPM (MPM) from healthy controls (HC) and from subjects without MPM but with a similar professional asbestos exposure (EXP) using noninvasive investigation. The identification of the pattern of compounds capable of diagnosing the disease was performed by the application of statistical treatments on the obtained data. The study was applied to patients with an established diagnosis of MPM, healthy controls and subjects with a certified long-term professional exposure to asbestos that had not developed the disease. The vision of the authors was to contribute to the development of breath analysis as an alternative noninvasive diagnostic method for the disease to blood and urine analysis.

Materials and methods

Subjects

A total number of 39 subjects volunteered to participate in this study. All the subjects were adults (45–80 years). The study population included three groups of 13 subjects: patients affected by MPM, subjects without MPM but with long-term professional exposure to asbestos and a healthy control group. Patients were recruited among those visiting the outpatient clinic of the Occupational Medicine Department, University of Bari, whilst controls were volunteers working in the hospital and in the university. Patients with a history of upper or lower respiratory tract infection during the past 4 weeks prior to the measurements and patients with asthma, chronic obstructive pulmonary disease, systemic diseases (such as diabetes) or a prior diagnosis of malignancy were not eligible for inclusion in this study. Current smokers were excluded from the study. The MPM group was composed of 13 patients with long-term professional asbestos exposure and a histology-established diagnosis of malignant pleural mesothelioma without current treatment by chemotherapy and/or radiotherapy. Patients were staged using the International Union Against Cancer TNM staging system [18]. The exposed group consisted of 13 subjects with long-term certified professional exposure to asbestos and with radiological signs of pleural plaques and/or benign asbestos pleural effusion. None of them suffered chest symptoms or other pulmonary and cardiovascular abnormalities. The healthy controls group also had 13 subjects each with a negative history of

professional exposure to asbestos, absence of any known diseases and no drug usage. The study was approved by the local Ethics Committee and all the subjects gave their written informed consent.

Study design

The study had a cross-sectional, case–control design with two visits within a 10-day period. The first day was a screening day to check all the inclusion and exclusion criteria. On the second day, exhaled breath was collected in duplicate and analysed. Subjects were asked to refrain from eating and drinking in the 3 h before the test.

Exhaled breath collection and sampling

The breathing manoeuvres were based on a validated method described in our previous paper [19]. In short, patients breathed tidally through a mouthpiece, connected to a three-way non-rebreathing valve and an inspiratory VOC-filter (A2, North Safety, NL) for 5 min. After a single deep inspiration, the patient exhaled a single vital capacity volume into a Tedlar bag connected to the expiratory port. These manoeuvres were done in duplicate by repeating the same procedure after a 2-min interval.

Analytical method

The analytical method of VOCs in exhaled breath consisted of three steps: adsorption of VOCs on sorbent cartridges, their TD and analysis by GC-MS [20]. The cartridge is composed of a cylindrical stainless steel net (100 mesh) with an external diameter of 4.8 mm, containing Carboxen 1003, Carbopack B, and carbopackY as adsorbent bed (Sigma Aldrich, Italy). The cartridge was connected on a side to the sampled Tedlar bag and on the other side to a low-flow sampling pump (Pocket Pump, SKC, Italy). For the sampling step, it was set to a flow of 25 mL/min for 30 min. In such conditions, it was aspirated a volume of 0.75 L per bag. Then, the sampled cartridges were thermally desorbed and the analytes were analysed by GC-MS. The analysis of VOC was carried out using a thermal desorber (Markes International Ltd., Unity™) equipped with an autosampler (Markes mod. ULTRA™ TD) provided with 100 positions and coupled with a gas chromatograph (Agilent GC-6890 PLUS) and a mass selective detector (Agilent MS-5973 N). The thermal desorber provides a two-stage mechanism: In the former, the analytes are desorbed from the sample tube and refocused into a cold trap; in the latter, they are desorbed from the trap and carried into the GC column. The operative conditions are listed in Table 1. The standard solutions were prepared by successive dilutions in methanol of a VOC standard mixture at 2,000 μg/mL containing all investigated analytes (Ultra Scientific Cus-5997). To quantify the samples, a calibration curve was prepared by injecting 1 μL of a standard solution into a tube; the spiked adsorbent tubes were then thermally desorbed in the same conditions of time, gas flow and split ratio of the samples.

Data statistical analysis

A variance analysis (ANOVA), principal component analysis (PCA) and discriminant function analysis (DFA) were

Table 1 Operative conditions for TD-GC-MS

Step	Parameter	Value
Adsorbent tube desorption	Purge time	1 min—trap in line
	Desorption time	10 min
	Desorption temperature	300 °C
	Temperature of cold trap	−10°C
	Desorption flow	20 mL/min
Focusing trap desorption	Desorption time	3 min
	Temperature of cold trap desorption	300 °C
	Split flow	44 mL/min
	Temperature transfer line	150 °C
GC analysis	Gas carrier	He
	Gas flow	1.7 mL/min
	Analytical column	SUPELCOWAX (Supelco), polyethylene glycol 30 m × 0.25 mm ID. 0.25 μm. Stationary phase thickness
	Oven temperature	40 °C for 3 min, 8 °C/min to 80 °C
		80 °C for 1 min, 20 °C/min to 270 °C
		270 °C for 3 min

applied to data. ANOVA is a statistical method used to compare the means of two or more samples. PCA is a multivariate analysis technique used to reduce data dimensionality to a smaller set of orthogonal factors of easier interpretation. PCA transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal factor), the second greatest variance on the second coordinate and so on. PCA involves the calculation of the eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix; the results of PCA are usually discussed in terms of component scores and loadings. DFA determines which variable allows discrimination between two or more groups. The method applies a multivariate *F* test in order to compare the total variances and covariance matrix with pooled within-group variances and covariances and to determine if significant differences concerning all variables occur between groups. The first discriminate function considered by the model provides the most discrimination between groups, the second provides the second most and so on. The basic assumptions of the procedure are: (1) normal distribution of the data; (2) the homogeneity in the variance/covariance matrices of variables across groups; and (3) the variables used to discriminate between groups are not completely redundant. The data were processed by means of the statistical software package Statistica for Windows, version 6.1.144.0 (StatSoft Italia srl, Vigonza, Italy).

Neural network

Counterpropagation artificial neural networks (CP-ANN) are one of the most popular neural networks proposed in literature and are increasing in uses and applications related to several multivariate chemical issues. CP-ANN is a modelling tool which combines features from both supervised and unsupervised learning [21, 22]. Moreover, a CP-ANN consists essentially of two layers, a Kohonen and an output layer [23]. In this paper, a graphical user interface of the Kohonen and CP-ANN toolbox in Matlab was used. The toolbox is a freely available collection of functions and algorithms and can be downloaded via the Internet from the Milano Chemometrics and QSAR Research Group web site (http://www.disat.unimib.it/chm).

Results and discussion

Setup of the analytical procedure

The first step of the present work was the setup of the analytical procedure starting from the method tested and used for environmental samples [20]. The preliminary analyses, conducted on healthy control samples, showed that the most abundant VOCs were: cyclopentane, phenol, 1,2-pentadiene, cyclohexane, heptane, methyl-cyclohexane, toluene, 2,4-dimethyl-heptane, methyl-octane, xylene, styrene, a-pinene, benzaldehyde, b-pinene, trimethylbenzene, decane, limonene, 2-ethyl-1-hexanol, *N*,*N*,dimethyl-acetamide (DMCA), acetophenone, dimethyl-nonane and dodecane. In order to ascertain the level of contamination in new or cleaned bags, background VOC concentrations in a clean Tedlar bag were monitored. Analysis of bags filled with clean and humidified (50%) air showed that DMCA and phenol are due to emissions of the bags themselves. It was known that these compounds are present in the production process of Tedlar [24]. Other studies also observed high backgrounds of DMCA and phenol in Tedlar bags [25, 26]. The cartridge desorption (first desorption) was performed in splitless mode in order to increase the sensitivity of the methodology. Limit of quantization (LOQ), expressed as nanograms per litre, was determined for all investigated compounds from the sequential analyses of diluted standard solution with a signal-to-noise ratio of 10 (Table 2).

Performance of the sampling and analytical method was evaluated using a VOCs standard mixture containing a known concentration of the target compounds. The sampling step requires the choice of the best adsorbent material and the optimisation of the flow rate and sampling time (sampling volume). A study of the different adsorbent materials showed that the triphasic adsorbent cartridges containing carboxen 1003, C2-C5, carbopack B, C5-C12, and carbopack Y, C12-C20, are the best place to absorb quantitatively the investigated compounds. Then, the flow rate and the sampling time (volume of air sampled) were tested and optimised using two cartridges connected in series to a bag containing a control sample. The best results were obtained in sampling under the following conditions: flow rate of 25 mL/min for 30 min. It was found that all investigated compounds were adsorbed quantitatively on the first cartridge; no compound was present in detectable quantities (<LOQ) in the second cartridge. Finally, method repeatability was evaluated analysing healthy control samples in triplicate. In particular, the sampling procedure was repeated three times for each HC individual and then the analysis of the three collected bags was performed. The relative standard deviation percentages (RSDs%) were <15% for all investigated compounds. About the discrimination of the gender (male/female), from the data analysis, it was noted that no significant variations were found linked to the different gender; the variability falls into that group owning.

Data analysis

The subject characteristics of the three groups are listed in Table 3: subjects with long-term exposure to asbestos,

Table 2 Medians of volatile compounds and RSD% for each investigated group

Volatile organic compounds	LOQ (ng/L)	HC		MPM		EXP	
		Median (ng/L)	RSD (%)	Median (ng/L)	RSD (%)	Median (ng/L)	RSD (%)
Cyclopentane	0.40	34.83a	96	120.42a	78	605.49b	77
1,2-Pentadiene	0.93	55.33a	70	64.20ab	85	200.66b	42
Cyclohexane	4.67	33.08a	58	251.79b	84	69.31c	130
Heptane	10.53	21.44a	139	22.08a	51	24.41a	48
Methyl-cyclohexane	8.13	18.12a	143	17.66a	47	30.33a	26
Toluene	0.93	12.99a	51	16.70a	36	15.37a	27
2.4-Dimethyl-heptane	8.40	14.02a	94	17.01ab	50	27.69b	32
Methyl-octane	8.40	14.06a	94	18.56a	55	32.70b	30
Xylene	2.27	13.50a	46	14.31ab	42	20.25b	33
Styrene	0.93	2.05a	138	4.09a	41	2.25a	104
a-Pinene	13.73	37.40a	53	36.51a	49	49.44a	26
Benzaldehyde	2.67	13.56a	53	14.41a	44	19.92a	23
b-Pinene	15.47	37.52a	51	35.52a	50	41.87a	33
Trimethylbenzene	2.67	12.34a	65	11.69a	48	16.18a	26
Decane	0.93	5.62a	50	6.59a	40	8.70a	27
Limonene	4.40	24.60a	151	23.73a	33	28.70a	49
2-Ethyl-1-hexanol	1.07	45.15a	42	48.24a	44	51.50a	28
Acetophenone	12.00	53.97a	40	51.61ab	50	70.30b	26
Dimethyl-nonane	15.47	16.65a	51	19.03a	45	36.71b	25
Dodecane	0.93	5.46b	34	7.49a	27	10.31a	21

LOQ, expressed as nanograms per liter, are listed in the second column

Results for three patient groups of HSD for unequal *N* Tukey test are reported: groups of one row with different letters are statistically different ($p<0.05$); volatile compound data were normalised before applying the test

patients affected by MPM, and healthy controls. Subjects with long-term exposure to asbestos were slightly older than healthy controls ($p<0.01$), whilst there were no significant differences in age between MPM and healthy controls as well as exposed subjects. The median values (in ng/L) and the correspondent RSD% for each compound in 39 subjects for the three groups are listed in Table 2. Median values were displayed because of the non-normal distribution of the raw data (Shapiro–Wilk tests). After a log transformation, a normal distribution of data into the groups was obtained for each compound ($W\geq0.9$), with the exception of heptane and 2-ethyl-1-hexanol ($W\leq0.8$). These last two variables were subsequently eliminated from statistical analysis.

Table 3 Clinical characteristics of the study population

	MPM	EXP	HC
Subjects (*n*)	13	13	13
Age (years)	60.9±12.2	67.2±9.8	52.2±16.2
Gender (M/F)	11:2	9:4	5:8
Ex-smokers (*n*)	5	4	5

A variance analysis (ANOVA), performed on normalised data, highlighted statistically significant differences ($p<0.05$) in the values of nine compounds among the patients belonging to three different groups. Despite these differences, examination of the results obtained applying a post hoc test (honestly significantly different (HSD) for unequal *N* Tukey) showed that cyclohexane was the only compound having mean values significantly different for all the three groups at the same time, whereas for other compounds, fewer (two or one) groupings are recognised (Table 2). In particular, MPM patients showed higher concentrations of cyclohexane (median value=251.79 ng/L) compared to subjects with long-term exposure to asbestos (median value=69.31 ng/L) and healthy controls (median value=33.08 ng/L). Cyclohexane is a known human metabolite of ε-caprolactam degradation [27], and this may support the hypothesis of a tight relationship between reactions of xenobiotic agents' degradation and the genesis of tumours since the relation between polluting agents and neoplastic processes is documented [28] (http://www.voc-research.at). Therefore, based on our results, cyclohexane seems to be associated with the presence of MPM and a possible biomarker for MPM disease.

ANOVA analysis also highlighted that the EXP averages for cyclopentane, methyl-octane and dimethyl-nonane were significantly different from the other two groups. This result shows that subjects without MPM but with long-term professional exposure are characterised by a breath composition different from both the MPM and HC groups. This could be a useful tool in the early diagnosis of MPM because it suggests the need of frequent and thorough checks on subjects with these characteristics.

PCA was used to provide a partial visualisation of data in a reduced-dimension plot. After applying PCA to our normalised data set, five PCs were extracted and the percentage of variance explained by each PC were 59.0%, 11.7%, 8.2%, 7.8% and 5.4%, respectively. Poor visual clustering was apparent when the scores of the patients were displayed with respect to the first two principal components (data not shown). This was not surprising since the first principal component accounts for the maximum possible one-dimensional projection of the total variation of the individual data points and each succeeding component accounts for the maximum possible one-dimensional projection of the remaining variability, which do not necessarily correspond to the maximum variations among defined classes. Therefore, with the aim of finding the best PCs able to discriminate among patients of the three different groups, an ANOVA was applied on the component score values. Statistically significant differences ($p<0.01$) were obtained only for PC4 and PC2 with F values of 24.4 and 17.1, respectively. In fact, when representing the scores of the patients on the two-dimensional space defined by these last PCs (Fig. 1), the patients seem to group into three clusters, corresponding to our different groups (HC, MPM, EXP), even if the separation is not completely satisfactory due to a partial overlapping of clusters. According to the loading of the variables (Table 4), the most contributing descriptors (loading value>0.50) were cyclopentane (−0.54) and 1,2-pentadiene (−0.83) for PC2 and cyclohexane (−0.73) and methyl-cyclohexane (0.55) for PC4.

Table 4 PCA loading components obtained on all investigated subjects

Variables	PC1	PC2	PC3	PC4
Cyclopentane	−0.68	−0.54		
1,2-Pentadiene		−0.83		−0.32
Cyclohexane	−0.40	0.23	0.46	−0.73
Methyl- cyclohexane	−0.71		0.27	0.55
Toluene	−0.88	0.23		
2,4-Dimethyl-heptane	−0.76	−0.26	0.35	0.40
Methyl-octane	−0.74	−0.43	0.37	0.27
Xylene	−0.95			
Styrene	−0.42	0.50	0.62	
a-Pinene	−0.92		−0.27	
Benzaldehyde	−0.96			
b-Pinene	−0.79		−0.45	
Trimethylbenzene	−0.91	0.25		
Decane	−0.89			
Limonene	−0.52	0.28	−0.32	
Acetophenone	−0.92		−0.26	
Dimethyl-nonane	−0.86	−0.43		
Dodecane	−0.82	−0.27		

Only component loadings with absolute values >0.2 are presented

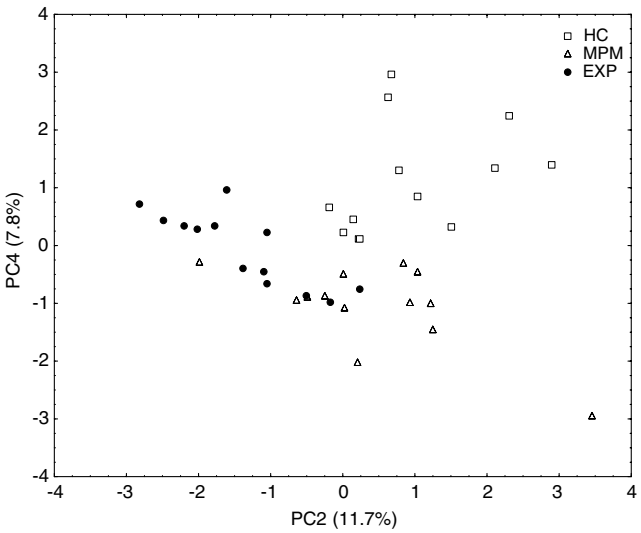

Fig. 1 Principal component analysis score plot of PC4 versus PC2 for all investigated subjects

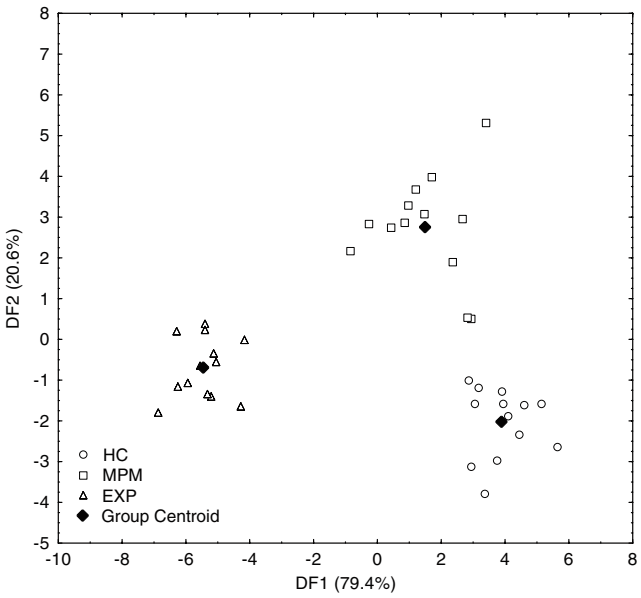

Fig. 2 A graphical representation of the discriminant function analysis (*DFA*) of the investigated subjects using volatile organic compound data

To assess the discrimination efficiency among the different patient groups, i.e. HC, MPM and EXP, a multivariate analysis of normalised data was carried out by DFA, and two different discriminant functions were computed with an accounted variability of 79.4% and 20.1%, respectively. As shown in Fig. 2, a complete separation of the three groups was achieved and the recognition ability was 100% for each group. The discrimination between EXP and the other two groups (MPM and HC) was clearly displayed along the first discriminant function (DF1), whilst the second function (DF2) led, to a minor extent, to the separation between all groups. In order to identify which compounds cause the discrimination, the factor structure matrix, in which the correlations between the variables and the discriminant functions were reported, was analysed. Cyclopentane (−0.35), 1,2-pentadiene (−0.17), dimethyl-nonane (−0.21) and dodecane (−0.18) were the dominant variables to define the first discriminant function whilst cyclohexane (0.46) and dodecane (0.17) for the second discriminant function. In particular, cyclopentane and cyclohexane were the most correlated variables with DF1 and DF2, respectively. This result is in agreement with that obtained by applying ANOVA. A leaving-one-out cross-validation procedure was used to evaluate the classification performance showing a prediction ability of 84.6% for the three groups. Among 39 subjects studied, five MPM patients were incorrectly classified—in particular two as HC and three as EXP—and one HC was misclassified as EXP. A stepwise DFA was also tested to select a subset of variables which, enabling the greatest separation between groups, discarded less significant or autocorrelated variables. Eleven variables were selected (cyclopentane, cyclohexane, dodecane, xylene, toluene, decane, methyl-cyclohexane, dimethyl-nonane, benzaldehyde, limonene, b-pinene), and even if the recognition ability was lower, i.e. 97.4% (one MPM patient was misclassified as HC), the cross-validation showed a higher prediction ability (94.9%) for the three groups. Only two MPM patients were incorrectly classified as HC. Moreover, cyclopentane (0.39) and cyclohexane (−0.49) were again the dominant variables to define the discriminant functions.

CP-ANN was also applied to data and networks settings were optimised in order to obtain the best performance in terms of classification. In particular, the following were determined: the network size (7), which defines the number of neurons for each side of the map; the number of epochs (600), which is the number of times each sample is introduced in the network; and the boundary condition (toroidal), which defines the network space. The learning rates were set by default at 0.5 and 0.01, respectively, as suggested in literature [21]. The fitting procedure using the optimised parameters showed a classification ability of 100% for each group. A leaving-one-out cross-validation procedure was used to evaluate the classification performance, showing a prediction ability of 87% for the three groups. Among the 39 subjects studied, five were incorrectly classified: two MPM patients were classified—one as HC and one as EXP—an HC was classified as MPM, and two EXPs were classified—one as HC and one as MPM.

Conclusions

The analytical method allowed the detection of the most abundant compounds in breath samples of healthy control, MPM patients and EXP subjects with high repeatability. Despite the small number of samples analysed, encouraging results were obtained. Univariate analysis performed on the data showed that cyclohexane is the only compound that differed significantly in the three investigated groups. The application of multivariate statistical treatments, PCA, DFA and CP-ANN, showed the importance of having the entire pattern of the compounds (fingerprint) to obtain a good discrimination between the groups (100%) and then to differentiate MPM patients from the HC and EXP groups. The application of stepwise DFA showed that using a small number of variables (cyclopentane, cyclohexane, dodecanoic, xylene, toluene, decane, methyl-cyclohexane, dimethyl-nonanoic, benzaldehyde, limonene, b-pinene), it was possible to obtain a recognition ability of 97.4%. All data treatments showed that cyclopentane and cyclohexane were the dominant variables able to discriminate between the three groups. In particular, it was found that cyclohexane is the only compound able to differentiate MPM group from the other two; therefore, it can be considered a possible marker of the investigated disease. Cyclopentane allowed discriminating EXP from the other two groups, so it can be considered a good indicator for long-term asbestos exposure. This result suggests the need of frequent and thorough checks on individuals without MPM but with a long-term professional asbestos exposure in order to constantly monitor their health. Therefore, it is possible to state that using breath analysis, it may be possible to diagnose the disease early using a noninvasive and highly accurate methodology.

References

1. Robinson BW, Lake RA (2005) Advances in previous malignant mesothelioma. N Engl J Med 353(15):1591–1603
2. Curran D, Sahmoud T, Therasse P, van Meerbeeck J, Postmus PE, Giaccone G (1998) Prognostic factors in patients with previous pleural mesothelioma: the European Organization for Research and Treatment of Cancer experience. J Clin Oncol 16(1):145–152
3. Grellier L, Astoul P (2008) Mesothelioma and asbestos-related plueral diseases. Respiration 76:1–15

4. Kaufman AJ, Pass HI (2008) Current concepts in malignant pleural mesothelioma. Expert Rev Anticancer Ther 8:293–303
5. Kent M, Rice D, Flores R (2008) Diagnosis, staging and surgical treatment of malignant pleural mesothelioma. Curr Treat Options Oncol 9:158–170
6. Greillier L, Baas P, Welch JJ, Hasan B, Passioukov A (2008) Biomarkers for malignant pleural mesothelioma. Mol Diagn Ther 12:375–390
7. Pauling L, Robinson AB, Teranishi R, Cary P (1971) Quantitative analysis of urine vapour and breath by gas–liquid partition chromatography. Proc Nat Acad Sci 68:2374–2376
8. Rooth G, Ostenson S (1966) Acetone in alveolar air and the control of diabetes. Lancet 2:1102–1105
9. Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, O'Neill HJ (1985) Volatile organic compounds in exhaled air from patients with lung cancer. Clin Chem 31(8):1278–1282
10. Phillips M (1997) Method for the collection and assay of volatile organic compounds in breath. Anal Biochem 247:272–278
11. Phillips M, Erickson GA, Sabas M, Smith JP, Greenberg J (1995) Volatile organic compounds in the breath of patients with schizophrenia. J Clin Pathol 48:466–469
12. Karl T, Prazeller P, Mayr D, Jordan A, Rieder J, Fall R, Lindinger W (2001) Human breath isoprene and its relation to blood cholesterol levels: new measurements and modeling. J Appl Physiol 91:762–770
13. Westhoff M, Litterst P, Freitag L, Urfer W, Bader S, Baumbach JI (2009) Ion mobility spectrometry for the detection of volatile organic compounds in exhaled breath of patients with lung cancer: results of a pilot study. Thorax 64:744–774
14. Pleil JD, Lindstrom AB (1997) Exhaled human breath measurement method for assessing exposure to halogenated volatile organic compounds. Clin Chem 43:723–730
15. Amann A, Smith D (2005) Breath analysis for clinical diagnosis and therapeutic monitoring. World Scientific Publishing, 5 Toh Tuck Link, Singapore
16. Belda Iniesta C, de Castro J, Carpeño CJA, Moreno V, Casado Sáenz E, Feliu J, Sereno M, García Río F, Barriuso J, González Barón M (2007) New screening method for lung cancer by detecting volatile organic compounds in breath. Clin Transl Oncol 9:364–368
17. Friedrich MJ (2009) Scientists seek to sniff out diseases: electronic "noses" may someday be diagnostic tools. JAMA 301:585–586
18. Sobin LH, Witterkind C (2002) TNM classification of malignant tumours. International Union Against Cancer, 6th edn. Wiley-Liss, New York
19. Dragonieri S, Schot R, Mertens JA, Le Cessie S, Gauw SA, Spanevello A, Resta O, Willard NP, Vink TJ, Rabe KF, Bel EH, Sterk PJ (2007) An electronic nose in the discrimination of patients with asthma and controls. J Allergy Clin Immunol 120 (4):856–862
20. Bruno P, Caselli M, de Gennaro G, Iacobellis S, Tutino M (2008) Monitoring of volatile organic compounds in non-residential indoor environments. Indoor Air 18:150–156
21. Zupan J, Gasteiger J (1999) Neural networks in chemistry and drug design, 2nd edn. Wiley-VCH, Weinheim
22. Zupan J, Novi M, Ruisánchez I (1997) Kohonen and counter-propagation artificial neural networks in analytical chemistry. Chemom Intell Lab Syst 38:1–23
23. Kohonen T (1988) Self-organization and associative memory. Springer, Berlin
24. Chase RE (2001) Properties and manufacture of Tedlar^{R} polyvinyl fluoride film VERL. Ford Research Laboratory ERC Technical Report Issued
25. Steeghs MML, Cristescu SM, Munnik P, Zanen P, Harren FJM (2007) The suitability of Tedlar bags for breath sampling in medical diagnostic research. Physiol Meas 28:73–84
26. Trabue SL, Anhalt J, Zahn JA (2006) Bias of Tedlar bags in the measurement of agricultural odorants. J Environ Qual 35:1668–1677
27. Gross P (1984) Biologic activity of epsilon-caprolactam. Crit Rev Toxicol 13:205–216
28. Landi MT, Consonni D, Rotunno M, Bergen AW, Goldstein AM (2008) Environment and Genetics in Lung Cancer Etiology (EAGLE) Study: an integrative population-based case–control study of lung cancer. BMC Public Health 8:203