Biomarkers, 2010; 15(4): 297–306

informa
healthcare

ORIGINAL ARTICLE

Non-invasive diagnosis of liver diseases by breath analysis using an optimized ion–molecule reaction-mass spectrometry approach: a pilot study

Gunda Millonig[1,2], Siegfried Praun[3], Michael Netzer[4], Christian Baumgartner[4], Albert Dornauer[3], Sebastian Mueller[2], Johannes Villinger[3], and Wolfgang Vogel[1]

[1]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Medical University of Innsbruck, Austria,[2]Center of Alcohol Research and Department of Internal Medicine, Salem Medical Center, University of Heidelberg, Heidelberg, Germany,[3]V&F Medical Development GmbH, Absam, Austria, and [4]Research Group for Clinical Bioinformatics, Institute of Electrical, Electronic and Bioengineering, University for Health Sciences, Medical Informatics and Technology (UMIT), Hall in Tirol, Austria

Abstract

Breath composition is altered in liver diseases. We tested if ion–molecule-reaction mass spectrometry (IMR-MS) combined with a new statistical modality improves the diagnostic accuracy of breath analysis in liver diseases. We analysed 114 molecules in the breath of 126 individuals (healthy controls, and patients with non-alcoholic and alcoholic fatty liver disease and liver cirrhosis) by IMR-MS. Characteristic exhalation patterns were identified for each group. Combining two to seven molecules in the new stacked feature ranking model reached a diagnostic accuracy (area under the curve) for individual liver diseases between 0.88 and 0.97. IMR-MS followed by sophisticated statistical analysis is a promising tool for liver diagnostics by breath analysis.

Keywords: *Diagnostics; mass spectrometry; foetor hepaticus; fatty liver; exhalation*

Introduction

Several major organs, such as colon, kidneys and liver are involved in the excretion of waste metabolites. In contrast, the excretory function of the lung has not gained adequate attention for diagnosis although many important volatile compounds are eliminated by ventilation, e.g. acetone in diabetic ketoacidosis. Because breath can be obtained non-invasively and its constituents directly reflect concentrations in blood, breath analysis appears to be an appealing method in the field of biomarker research. Indeed, analysis of the human breath with mass spectrometry (MS) has been reported to provide useful information in patients with lung cancer (reviewed in Mazzone 2008) and asthma (Kharitonov et al. 1994), and to estimate the level of oxidative stress end products (Risby & Sehnert 1999) or previous exposure to potentially harmful xenobiotics (Periago et al. 1994).

The liver as the central metabolic organ is the source of many volatile organic compounds (VOCs) that are exhaled, e.g. ketones or sulphur-containing components such as dimethyl sulphide (Van den Velde et al. 2008) or carbonyl sulphide (Sehnert et al. 2002). Changes in breath composition in advanced liver disease can be readily detected with the human olfactory system and the so-called 'foetor hepaticus' has had an important diagnostic value for many centuries. Thus, techniques such as MS should be able to define minimal changes in breath composition

The first two authors contributed equally to this work.

Address for Correspondence: Gunda Millonig, Center of Alcohol Research and Department of Medicine, Salem Medical Center, University of Heidelberg, Zeppelinstrasse 11-33, 69121 Heidelberg, Germany. Tel: +49-6221-483-201. Fax: +49-6221-483-494. E-mail: gunda.millonig@uni-heidelberg.de

(Received 25 October 2009; revised 12 January 2010; accepted 13 January 2010)

ISSN 1354-750X print/ISSN 1366-5804 online © 2010 Informa UK Ltd
DOI: 10.3109/13547501003624512

http://www.informahealthcare.com/bmk

even in earlier stages of liver disease quantitatively. Indeed, isolated VOCs could be identified in liver diseased patients using gas chromatography followed by MS (GC-MS) (Mazzanti et al. 1989, Sehnert et al. 2002, Tangerman et al. 1983, Van den Velde et al. 2008). Although the results elucidated some of the differences in breath between patients with liver diseases and healthy controls, the GC-MS approach had two major limitations. First, the spectrum of analytes was rather narrow. Second, additional preconcentration steps were required which are cumbersome and give rise to potential biases.

In this article, we present a pilot study of human breath analysis in liver disease using an optimized MS device with regard to sampling, sample processing and ionization techniques. Based on several selected breath components, we finally present novel diagnostic algorithms to discriminate between various different liver diseases.

Material and methods

Patient selection

All patients gave their informed consent to breath analyses. Tests were approved by the guidelines of the local ethical committee and performed according to the Helsinki declaration. Patients were selected according to their diagnosis and divided into three groups: non-alcoholic fatty liver disease (NAFLD; $n = 34$, age 49.0 ± 14.6 years), alcoholic fatty liver disease (AFLD; $n = 20$, age 49.5 ± 12.2 years) and liver cirrhosis (viral or alcohol-related, $n = 37$, age 54.8 ± 11.6 years). Diagnosis of liver diseases was established by ultrasound, blood tests and clinical history. For ethical reasons, patients were not characterized by liver biopsy. Patients' characteristics are given in Table 1.

The control cohort consisted of 35 healthy individuals (mean age 37.4 ± 14.3 years) with no history of liver disease and normal lab tests including transaminases, liver synthesis parameters, glucose and iron metabolism and viral hepatitis markers. Routine blood tests included aspartate aminotransferase, alanine aminotransferase, γ-glutamyl transferase, alkaline phosphatase, C-reactive protein (CRP), creatinine and urea. Individuals with advanced steatosis or cirrhosis were additionally tested for prothrombin time and serum albumin. Exclusion criteria for the study were a calculated creatinine clearance < 60 ml min^{-1}, elevated CRP levels (>0.7 mg dl^{-1}) or a transjugular intrahepatic portosystemic shunt (TIPS). All groups of patients comprised a similar percentage of smokers and non-smokers. In the liver cirrhosis group there were no patients with higher degree encephalopathy.

Breath sampling

Individuals had to fast and abstain from smoking overnight and they also had to be at rest for at least 15 min prior to the measurements. Whole breath samples were collected by exhalation through a drinking straw dipped into a small glass vial of 20 ml volume (Macherey-Nagel, Düren, Germany). Patients were then asked to exhale once through this drinking straw into the vial. After completed exhalation, glass vials were crimped airtight and transferred for immediate analysis or frozen at -20°C until analysed. All samples were taken in duplicate to guarantee reproducibility.

Quality control of breath samples

As exhaled air is enriched in CO_2, this gas can serve as a quality marker to exclude air-contaminated samples from the study. Therefore, the concentration of CO_2 was assessed in all samples after MS analysis. A CO_2 concentration in the samples higher than 0.4 vol% during the whole analysis period was considered non-contaminated. Samples that did not meet this requirement were excluded from further analysis. In order to guarantee stability of the samples, they were either analysed immediately after collection or frozen at -20°C, which has been shown to keep breath compounds stable for up to 3 weeks (Bennett et al. 2009).

In a small-scale preliminary study we collected ten breath samples per person and determined reproducibility. Intraindividual coefficient of variation (CV) for molecules with concentrations above 1 ppb ranged from 3.09% to 70.05% with a mean CV of 16.7%. A CV of over

Table 1. Patients'characteristics.

Group	No. of patients	Sex (M/F), *n*	Age (years)	BMI (kg m^{-2})	AST (U l^{-1}) normal (<18 U l^{-1})	ALT (U l^{-1}) normal (<22 U l^{-1})	GGT (U l^{-1}) normal (<28 U l^{-1})	AP (U l^{-1}) normal (60–170 U l^{-1})
Healthy	35	17/18	37.4 ± 14.3	22.2 ± 3.0	10.0 ± 2.2	11.1 ± 3.7	11.1 ± 4.1	81.3 ± 20.7
NAFLD	34	25/13	49.0 ± 14.6	26.0 ± 3.2	15.0 ± 5.7	31.1 ± 18.7	60.7 ± 31.1	121.3 ± 33.4
AFLD	20	16/4	49.5 ± 12.2	23.9 ± 2.5	25.9 ± 15.5	27.8 ± 15.8	124.3 ± 148.6	125.4 ± 55.0
Cirrhosis	37	25/12	54.8 ± 11.6	25.6 ± 5.3	30.2 ± 18.4	24.9 ± 15.9	74.4 ± 66.7	147.8 ± 54.4

BMI, body mass index; AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, γ-glutamyl transferase; AP, alkaline phosphatase; NAFLD, non-alcoholic fatty liver disease; AFLD, alcoholic fatty liver disease.

40% was arbitrarily set as the exclusion criterion for further analysis. Nine masses (7% of detected masses) were therefore excluded. We then noticed that even with two breath samples we had comparable reproducibility rates and therefore breath samples were collected in duplicate throughout the whole study.

Negative control

In order to prevent interference of the patient's breath with contaminating substances in the ambient air, one additional vial was sampled with ambient air of the room where the sampling procedure took place and then compared with the breath samples. After ion–molecule reaction (IMR)-MS analysis of breath samples and ambient room air, we calculated the 'positive alveolar gradient' for each mass, i.e. concentration in exhaled breath minus concentration in ambient room air.

Mass spectrometry analysis

The V&F Airsense.net (V&F, Absam, Austria) is an IMR mass spectrometer (Futrell 1986). The IMR-MS system used in this study was originally designed to measure trace gas components in industrial fields such as fuel cell or reformer gas development, work space control and environmental measurements (Airsense Mass Spectrometry Systems; V&F Medical Development, Absam, Austria). The system is based on the use of ion–molecule reactions coupled with quadrupole MS and provides a highly sensitive method for online and offline sampling of organic and inorganic compounds in exhaled breath. The use of Airsense.net for breath analysis has been recently described by Hornuss and colleagues (2007).

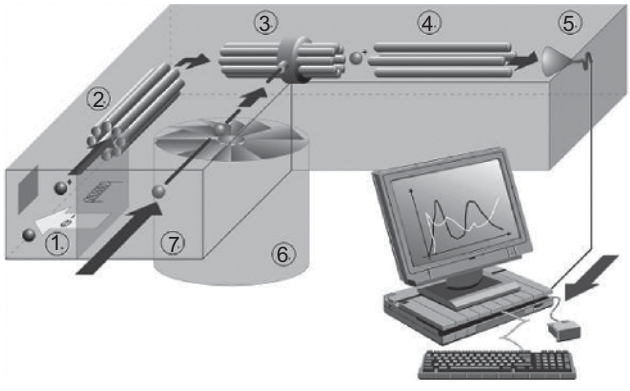

Figure 1. Principle of ion–molecule reaction mass spectrometry (IMR-MS) (for a detailed description see Material and methods). Important sections of the device are the following: 1, primary ion source; 2, octopole separation device; 3, charge exchange cell; 4, quadrupole mass filter; 5, particle detector; 6, vacuum system; 7, gas inlet system.

A schematic diagram of the IMR-MS is shown in Figure 1. In the IMR-MS analyzers, the principle of ion–molecule reactions is applied as the interaction of positively charged atomic ions with neutral sample gas molecules in two body collision processes resulting in the formation of product ions whenever the ionization potential of the sample molecule is less than the potential energy of the incoming primary ion and hence the entropy of the process becomes positive. The excess energy of the binary reaction is first stored in the product ion as transition state and either is statistically distributed in internal degrees of freedoms (electron vibration, bond oscillations) or is used up to break the weakest bond of the ionized molecule leaving a lower-molecular-weight ion.

Differences in ionization potentials between primary and product ions may result in a bond rupture and hence a lower-molecular-weight fragment ion. The IMR-MS uses krypton, xenon or mercury atomic gas to form the primary ion beam via electron impact ionization (section 1, Figure 1). The patented IMR ionization method can use the atomic mass scale to detect different molecules with the same molecular weight. As an example, acetaldehyde and CO_2 have the same mass (44). The mercury beam with an ionization potential of 10.4 eV does not ionize carbon dioxide (13.8 eV), but does ionize acetaldehyde (10.2 eV). Switching different ion beams and hence energy levels is fast and takes 400 ms. Krypton ions (14.0 eV) separate nitrogen (15.6 eV – not ionized) against carbon monoxide (13.7 eV) on mass 28.

The instrument uses two octopole systems (sections 2 and 3, Figure 1) operated at high frequencies to store primary as well as product ions in a confined volume against their coulomb repulsion and transmit ions to the quadrupole mass analysing section. The quadrupole mass separator (section 4, Figure 1), driven by direct current and alternating current, operates as an electromagnetic filter according to a parametric resonance to a specific mass over charge ratio. At a given alternating to direct current ratio, only one specific mass of ions experiences a stabile trajectory through the quadrupole. A secondary electron multiplier (section 5, Figure 1) may generate as much as 108 electrons for each incoming ion. This allows the generation of an electrical pulse strong enough to be accepted by a computer counting system. The pulse rate represents the concentration of the molecular species in the gas sample brought to the instrument. The sample gas (section 7, Figure 1) is transferred in a 2.5-m-long heated capillary system (Silcosteel®; Restek, Bellefonte, PA, USA) at a flow rate of 50 ml min^{-1} to the instrument. A constant pressure controller feeds via a second capillary a stable amount of 1.5 ml min^{-1} into the high vacuum ionization section.

Statistical analysis

Data are expressed as mean ± SD. Independent continuous variables were compared by means of the independent two-sample *t*-test and Mann-Whitney *U* test, and three additional feature selection methods (details see below). For comparisons of more than two variables, the Kruskal–Wallis test (H-test) with Bonferroni's correction was performed. The level of statistical significance was set to $p < 0.05$.

For the detailed biostatistic approach we refer to the recently published paper by Netzer et al. (2009). In brief, a new type of feature selection was performed. Feature selection aims at reducing the number of parameters to a smaller generalizing subset with high discriminatory ability. This technique ranks the features (i.e. molecular masses in breath) according to their quality to discriminate between predefined classes (i.e. groups of liver diseases) in a dichotomous way. A reduction of the number of features is then achieved by choosing a set of *k* best-ranked features. In our case this has been done by stack feature ranking (SFR). SFR is the combination of four different feature ranking techniques, BioMarkerIdentifier (Baumgartner & Baumgartner 2006), Information Gain (Quinlan 1993), ReliefF (Kononenko 1994), and using the *p*-value from a statistical significance test generating a consensus ranking of analysed breath gas molecules. The molecular mass with the highest discriminatory ability was identified, added to the algorithm and recalculation for the next best feature was started until addition of any other feature did not improve the discrimination power of the algorithm any further. The discriminatory ability of selected molecule subsets is represented by the area under the receiver operator characteristic curve (AUC).

Finally, the predicted best features were validated by cross-validation. For our experiments we used a stratified tenfold cross-validation strategy. In this type of cross-validation, the original sample is randomly partitioned into ten subsamples. Of the ten subsamples, a single subsample is retained as the validation data for testing the model, and the remaining nine subsamples are used as training data. The cross-validation process is then repeated ten times (the folds), with each of the ten subsamples used exactly once as the validation data. The ten results from the folds can then be averaged to produce a single estimation. The advantage of this method over repeated random subsampling is that all observations are used for both training and validation, and each observation is used for validation exactly once.

We compared single liver diseases among them and with healthy controls. Finally we also analysed the whole cohort with liver diseases compared with healthy controls.

Results

Soft ionization MS allows a broad characterization of human breath

We measured 114 molecular masses in human breath. Seventy-eight of these masses had a positive alveolar gradient (i.e. they were present in the exhaled breath but not in the surrounding air control vial). The masses were tentatively identified using V&F–ACP software. Identified masses ($n = 28$) are presented with their trivial name, the others remain to be identified and are referred to as molecular masses (e.g. molecular mass 76 is denoted as M76).

Different masses showed different intraindividual reproducibility. Generally, masses that were exhaled abundantly showed more constant concentration in breath samples (e.g. methane, N_2O, M39). Masses close to the lower detection limit showed higher variability (e.g. M40, M50, M117).

Different liver diseases are reflected by different breath patterns

Forty-four molecular masses out of the 114 measured showed a differential exhalation pattern in the four liver disease cohorts, 19 of them significantly at the $p < 0.05$ level (Table 2). We measured the most prominent differences between patients with AFLD, NAFLD, liver cirrhosis and healthy individuals in the biomarkers acetaldehyde, endogenous ethanol, isoprene, methane and hydrogen sulphide.

Table 2. Molecular masses in human breath showing statistically significant differences in the four groups analysed.

Molecular mass	*p*-Value(Kruskal–Wallis test)
Methane	0.0002
Dinitrogen monoxide	0.0002
Nitrogen monoxide	0.0001
Methylamine	0.0001
Molecular mass 32	<0.0001
Hydrogen sulphide	<0.0001
Propene	0.0003
Acetaldehyde	0.0001
Ethanol	<0.0001
Molecular mass 49	0.0002
Butadiene	0.0004
Molecular mass 60	<0.0001
Molecular mass 67	0.0001
Isoprene	0.0001
Molecular mass 76	<0.0001
Benzene	<0.0001
Molecular mass 79	<0.0001
Molecular mass 100	0.0001
Molecular mass 103	<0.0001

While acetaldehyde was comparable in healthy individuals and cirrhotics, it was significantly increased in patients with fatty liver disease, independent of alcoholic or non-alcoholic origin (Figure 2A). Breath ethanol increased only in patients with liver cirrhosis but not in those with AFLD or NAFLD compared with healthy controls (Figure 2B). Isoprene was significantly increased in patients with AFLD compared with all other groups (Figure 2C). Among masses without known trivial names, M103, M76, M39, M80 and M75 showed the highest differences.

Data analysis using a 'stack feature ranking' model increases discrimination ability between different liver diseases

As the difference in breath compounds between the different liver disease conditions is not associated with the appearance of specific single aberrant metabolites, an algorithm is needed that takes into account that there is a quantitative difference in several molecules (a 'breath fingerprint'). For this purpose we used a newly developed SFR model and compared it with established models. Compared with previously used feature ranking models the newly designed SFR model showed higher discrimination ability. Figure 3 shows that SFR is better than the single feature selection methods 'Information Gain' (Quinlan 1993), 'ReliefF' (Kononenko 1994), 'BioMarkerIdentifier' (Baumgartner & Baumgartner 2006) and using the *p*-value from a statistical significance test with a 10–15% higher AUC when distinguishing between AFLD and healthy controls. The algorithm selected a subset of five breath gas molecules, achieving the maximal AUC of 0.97 with a sensitivity of 0.95 and an optimal specificity of 1.0. Figure 4A–C shows the predictive values of models discriminating best between the different groups of liver disease. While patients with AFLD differ from healthy controls with an AUC greater than 0.9 after combining only two ranked molecular masses (Figure 4A), six ranked masses are needed to reach the same discriminatory precision between NAFLD patients and healthy controls (Figure 4B) and

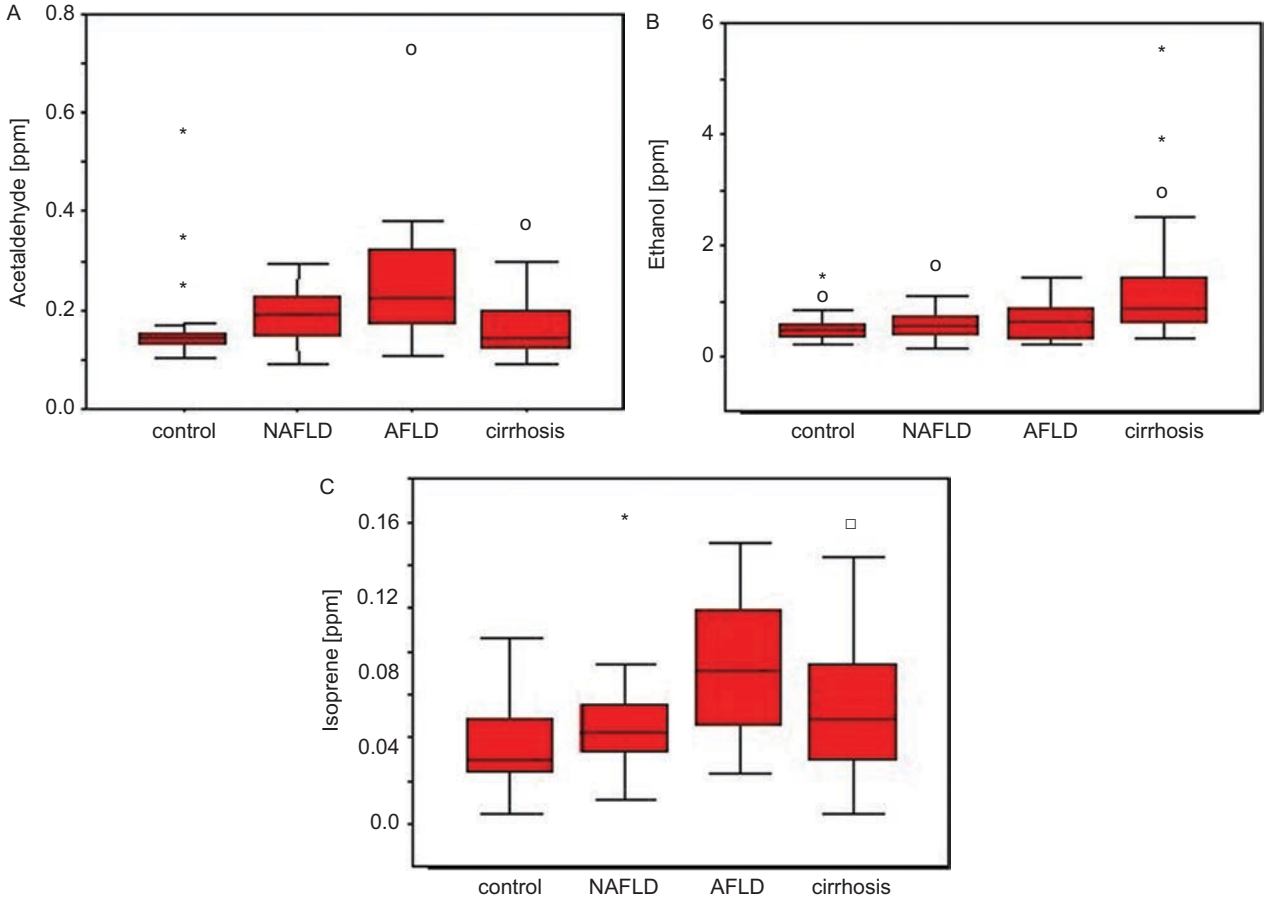

Figure 2. Three isolated molecules in liver disease breath analysis with a distinct pattern for different liver diseases. Data are presented as box plots. This type of graph represents lower, median and upper quartile in form of the box. Sample minimum and maximum are represented as whiskers. (A) Increased levels of acetaldehyde occur in patients with non-alcoholic fatty liver disease (NAFLD) and alcoholic fatty liver disease (AFLD) compared with healthy people and liver cirrhosis patients. (B) Endogenous ethanol levels increase in patients with liver cirrhosis. (C) Isoprene exhalation is increased in all liver disease groups compared with the healthy control cohort.

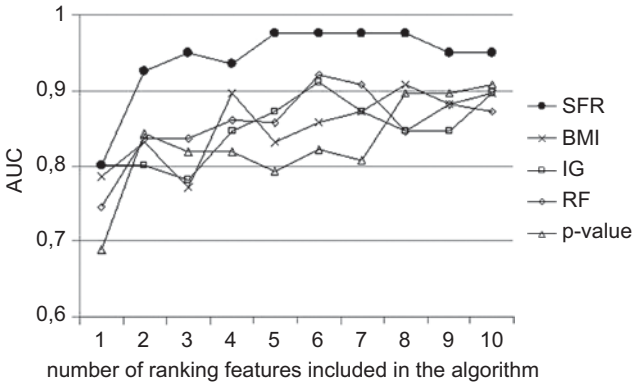

Figure 3. Area under the curve (AUC) of the ten molecules best discriminating between alcoholic fatty liver disease (AFLD) and healthy controls, ranked by the single methods Information Gain (IG), ReliefF (RF), BioMarkerIdentifier (BMI) and null-hypothesis testing (*p*-value) versus stack feature ranking (SFR). The discrimination between patients with AFLD and healthy controls is chosen as a representative example.

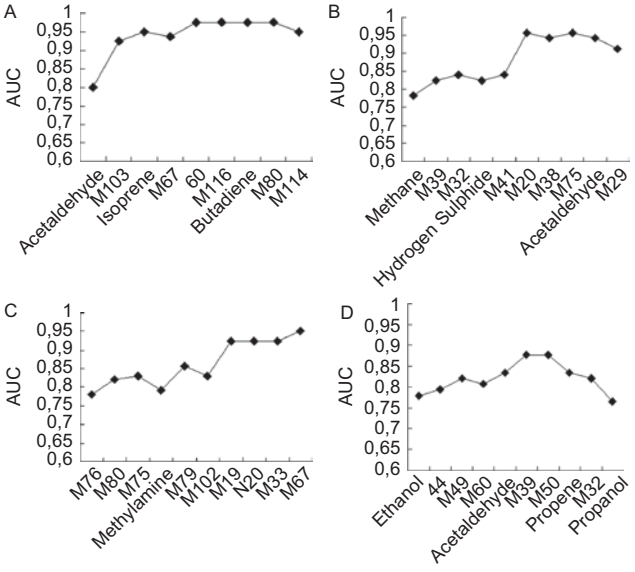

Figure 4. Stack feature ranking (SFR) analysis allows models to be calculated that take into account up to the ten top-ranked breath molecules for each patient group. (A) This shows the discrimination ability between alcoholic fatty liver disease (AFLD) patients and healthy individuals (best area under the curve (AUC) = 0.97); (B) non-alcoholic fatty liver disease (NAFLD) patients versus healthy individuals (AUC = 0.96); (C) NAFLD versus AFLD patients (AUC = 0.95); (D) between patients with liver cirrhosis and healthy individuals (AUC = 0.88).

seven ranked masses to discriminate AFLD from NAFLD (Figure 4C) patients by breath analysis. The combination of the six best-ranked molecular masses to differentiate cirrhotic patients from healthy individually increases the AUC to a maximum of 0.88 (Figure 4D). Numbers of maximal AUC values, standard deviations and the 95% confidence interval are given in Table 3. The experiments show that SFR is able to identify a small subset of five to seven molecular masses in this dataset with higher discriminatory and predictive value compared with traditional null-hypothesis tests and other biostatistical methods.

Table 3. Maximal areas under the curve from the tenfold cross-validation.

Compared groups	Mean	SD	95% Confidence interval
AFLD vs healthy	0.97	0.079	0.926–1
NAFLD vs healthy	0.96	0.074	0.907–1
AFLD vs NAFLD	0.95	0.105	0.907–1
Liver cirrhosis vs healthy	0.88	0.101	0.817–0.942
Liver disease vs healthy	0.94	0.080	0.883–0.981

AFLD, alcoholic fatty liver disease; NAFLD, non-alcoholic fatty liver disease.

Influence of smoking on breath analysis

As smoking strongly influences breath composition, we tested if smoking interferes with breath analysis for liver diagnostics. Within our healthy control cohort, 31.4% were smokers and 68.6% were non-smokers. We found that a total of 46 molecular masses were significantly different in the breath of smokers compared with those of non-smokers. Among the annotated molecules we found acetonitrile, formaldehyde, sulphur dioxide, methanol, propanol and acetic acid. None of the masses (annotated or unannotated) that differed significantly between smokers and non-smokers was found overlapping with the discriminating masses for liver diseases.

Discussion

Human breath contains a representative mixture of volatile and soluble compounds generated during metabolic processes in the body. As the liver is the main central organ in metabolism, we investigated whether breath analysis using IMR-MS is a feasible approach for liver diagnostics in the clinical setting.

The sampling procedure of collecting breath in a glass vial with a crimp cap proved to be easy to handle, reliable for storage of the breath sample and reproducible for transfer of the breath sample into the mass spectrometer. IMR-MS does not need any preconcentration step before analysis compared with other methods used for breath analysis in liver diseases (Sehnert et al. 2002, Van den Velde et al. 2008) and it proved to be a sensitive method for measuring concentrations as low as a few ppt (parts per trillion). At the same time its wide detection range allows analysis of the more abundant masses, e.g. CO_2, H_2O and O_2. The absence of fragmentation allows analysis of complex gas mixtures (an overview of the technical advantages is given by Tegtmeyer et al. (1993)).

Detection of CO_2 as a marker for hermetic sealing of the breath sample proved to be a good tool for assessing quality of the measured sample.

Standardization of breath analysis is a difficult matter. So far it is has been achieved only in breath analysis of single molecules such as NO (Bohadana et al. 2008, Takalo et al. 2008) but to the best of our knowledge not for multimass detection systems such as IMR-MS. In our study we also attempted to investigate the reproducibility of breath analysis by sampling two vials from each individual. Reproducibility in two vials was acceptable with a median CV of 16.7% between the vials for masses exhaled in greater concentrations than 1 ppb. We are aware that this is far from perfect. However, this variation is more due to the 'human factor', i.e. variation in low abundant masses from breath to breath as the detection limit of Airsense.net has been identified in the low ppt range and linearity has been proven down to low ppb concentration of breath components (Dolch et al. 2008). Unfortunately other groups measuring more than one molecule at the same time in human breath have not commented on reproducibility or compared standardized gas samples versus human breath.

Dietary influence on breath analysis is also a matter of concern and is even more difficult to standardize. From H_2 breath tests routinely used for diagnosis of lactose or fructose malabsorption we know that reliable results are achieved after an overnight fast. However, for direct influence of different diets on breath composition, detailed studies of individuals on a standard diet are needed.

Another possible interference in breath analysis is smoking. It has been shown in several publications that breath analysis allows differentiation between smokers and non-smokers. Among the molecules known to be present at increased concentrations in the breath of smokers are aromatics, e.g. benzene (Wallace et al. 1987), and also other components such as acetonitrile (Buszewski et al. 2009). While benzene is only a marker of recent smoking (within 1 h after the last cigarette it decreases to non-smoker levels), acetonitrile has a longer half-life time and can be detected up to 1 week after the last cigarette (Jordan et al. 1995). In order to test if smoking interferes with breath analysis for liver disease, we performed a subgroup analysis between smokers and non-smokers in the cohort without liver diseases. We found a significant difference in 46 exhaled masses. We could confirm the significant increase in acetonitrile but also found increased levels of formaldehyde, methanol, sulphur dioxide, propanol and acetic acid. As expected, there was no difference in benzene levels because individuals were asked to abstain from smoking overnight. When comparing the list of masses differentiating between smokers and non-smokers, there was no overlap with those differentiating between any of the liver disease groups. These results show that the differences in exhaled breath caused by smoking were abolished by selecting a similar percentage of smokers in all four groups analysed. It was encouraging to learn that different coexisting pathophysiological conditions might not necessarily preclude the diagnostic use of breath analysis when care is taken to exclude influence on the same molecules.

Using IMR-MS we detected 114 compounds in human breath. This number appears low in contrast to previous publications where authors have measured more than 1000 compounds by MS (Phillips et al. 1999). The reason for this discrepancy lies in the technical approach. While conventional MS leads to a high degree of fragmentation and thus disassembles each molecule to several peaks, this is not the case in IMR-MS. It therefore seems reasonable that these 114 masses will basically represent the same spectrum of breath compounds as the higher number published by Phillips and colleagues (1999).

The ideal breath test would identify specific single aberrant metabolites that are characteristic for specific liver diseases. This is not the case as our data and previously published data have proven. Breath analysis showed that differences in patterns of compounds between healthy controls and patients with AFLD, NAFLD and liver cirrhosis are quantitative more than qualitative, i.e. there is no compound that can be used as an exclusive marker for any of the diseases. In this series of patients, all 114 masses were present in each individual.

Sophisticated statistical modelling is therefore a useful tool in order to extract the most information from breath analysis. This was the reason for us to use a new ensemble feature selection modality based on the combination of different existing algorithms embedded in a computational workflow including data import, data preprocessing and a breath gas biomarker search. In this process, the molecules with highest discriminatory power were selected for each liver disease and then combined stepwise to increase sensitivity and specificity for diagnosis of the respective liver disease. With AUC results of 0.88 (liver cirrhosis vs healthy control), 0.94 (liver disease vs healthy control), 0.96 (NAFLD vs healthy), 0.97 (AFLD vs healthy), and 0.95 (AFLD vs NAFLD) using a tenfold cross-validation strategy we were able to improve sensitivity and specificity compared with conventional discrimination analyses.

It is surprising that the discrimination between healthy controls and liver cirrhosis is lower than between other groups of diseases. One explanation might be the inclusion of patients with alcoholic and viral liver cirrhosis. When we extended the SFR model to a combination of blood tests and breath compounds the discrimination ability to distinguish between healthy and cirrhotic individuals the AUC can be increased to 0.92 (vs 0.88). Combining conventional blood tests with

breath compounds might be an interesting option for breath analysis.

Despite the fact that there is no absolute marker for liver disease, some of the molecules detected deserve a closer look, because they can be linked to liver physiology or pathology.

Some of the molecules that show the highest difference between the four groups, such as isoprene, have been associated with liver disease in previous publications. Isoprene is one of the best characterized components in human breath. Our results show increased isoprene concentrations in AFLD compared with the other three groups. Others have shown that breath isoprene is linked to cholesterol synthesis as a byproduct derived from mevalonate (Stone et al. 1993). It might therefore correlate with a disturbed cholesterol biosynthesis in alcoholic liver disease.

Another interesting aspect in breath analysis is endogenous ethanol. Regarding endogenous ethanol levels in breath, cirrhotic individuals demonstrate increased concentrations compared with all the other groups. An increase of endogenous ethanol in cirrhotic patients might be due to either increased shunting volumes through portocaval shunts which prevents metabolism of endogenous ethanol in the liver. Diminished ethanol oxidation would also explain the decreased concentration of acetaldehyde in patients with liver cirrhosis compared with patients with alcoholic or non-alcoholic fatty liver.

The phenomenon of basic ethanol levels in breath was described many years ago (Jones 1985) and is thought to arise from normal metabolism (Jones et al. 1984) although it has been reported that bacterial overgrowth can increase endogenous ethanol levels drastically (Spinucci et al. 2006). Rat experiments have clearly shown that the intestinal microenvironment has an impact on ethanol metabolism (Baraona et al. 1986). This finding is emphasized by the fact that experimental modifications of the intestinal flora by feeding lactobacillus (Nanji et al. 1994) as well as by antibiotic treatment (Adachi et al. 1995) change the course of liver disease. In morbidly obese patients with non-alcoholic steatohepatitis an increased amount of ethanol has been measured (Solga et al. 2006). Others have shown that obesity as well as female gender is associated with breath ethanol concentrations (Nair et al. 2001). Obese mice, an animal model for NAFLD also show increased breath ethanol, underscoring the role of endogenous ethanol production in fatty liver pathogenesis (Cope et al. 2000). All of the above-mentioned observations make the interplay between intestinal flora, ethanol production and ethanol metabolism interesting with regard to a common pathogenetic pathway in AFLD and NAFLD.

The most interesting isolated molecule for the use of a breath test in liver diseases might be acetaldehyde. Acetaldehyde was present at increased concentrations in patients with AFLD or NAFLD but not in cirrhotics or healthy people. Acetaldehyde is the first oxidation product in ethanol metabolism. Under normal circumstances it is mainly generated by alcohol dehydrogenase and only in traces by cytochrome p450 2E1 (Cyp2E1) and catalase. In chronic alcohol abuse (Takahashi et al. 1993), but also in non-alcoholic fatty liver (Weltman et al. 1998) probably due to ketones and fatty acids, Cyp2E1 is induced and leads to increased acetaldehyde levels. Increased acetaldehyde levels in turn promote lipid peroxidation, liver cell damage but also carcinogenic etheno-DNA adducts (reviewed in Seitz & Stickel 2007)). Cyp2E1 induction is strongly associated with lipid peroxidation and accumulation of carcinogenic ethno-DNA adducts (Wang et al. 2009) but shows pronounced interindividual variability after induction with for example ethanol (Oneta et al. 2002). So far there has been no predictive marker to assess whether a patient with fatty liver is at risk for damage due to Cyp2E1 induction. Testing for acetaldehyde concentrations in the breath might be a first approach to risk stratification, but prospective studies are still needed.

Overall we can conclude from our pilot study that breath analysis by IMR-MS is easy to handle, feasible in the clinical setting, and allows adequate quality control of the breath samples for ambient air components and vial leakage. Breath analyses in liver disease patients yielded interesting results with regard to individual molecules, e.g. acetaldehydes, that might be worth following up in further prospective trials. Finally, we show that the new ensemble-based algorithm is a useful tool for identification of biomarker patterns.

Nevertheless, the problematic part of breath analysis lies deeper and is most probably independent of the applied measuring technique. One of the major difficulties in breath analysis is interference from basic factors such as nutrition but also from coexisting pathological conditions, e.g. chronic obstructive pulmonary disease and liver cirrhosis. We certainly need to study patient cohorts with isolated diseases and compare them with patients with more than one disease to learn about these interferences. Another issue is reproducibility and standardization of breath sampling; before breath analysis is ready for clinical routine, questions such as the need of standardized meals the day before sampling must be clarified. Also if we want to move forward from using breath analytes as biomarkers we also need to increase our knowledge on how volatile organic and inorganic compounds are produced by hepatic metabolism.

Acknowledgements

We thank Gisela Egg and Dr Margit Egg for their help with breath sampling, Julia Rössler-Graf for performing

analysis and documentation and Dr Thomas Walch from the Department of Psychiatry for his help recruiting the patients. This project was supported by the 'Kompetenzzentrum Medizin Tirol' (KMT 09), the 'Verein zur Förderung wissenschaftlicher Arbeit in Gastroenterologie und Hepatologie in Innsbruck', and the Austrian GEN-AU project 'Bioinformatics Integration Network'. G.M. is a recipient of an Olympia-Morata-Fellowship from the Medical School of Heidelberg.

Declaration of interest

J.V., S.P. and A.D. have financial interests regarding the IMR-MS system. None of the other authors have received funding or other financial support from V&F Medical Development; this includes contracts, equity interest, stock options, direct or indirect salary support, consultant fees, lecture/travel fees, or honoraria received within a period of 5 years of the date of submission of this manuscript. The funding sources had no role in the study design, in the collection, analysis, and interpretation of data, or in the decision to submit the manuscript for publication.

References

Adachi Y, Moore LE, Bradford BU, Gao W, Thurman RG. (1995). Antibiotics prevent liver injury in rats following long-term exposure to ethanol. *Gastroenterology* 108:218–24.

Baraona E, Julkunen R, Tannenbaum L, Lieber CS. (1986). Role of intestinal bacterial overgrowth in ethanol production and metabolism in rats. *Gastroenterology* 90:103–10.

Baumgartner C, Baumgartner D. (2006). Biomarker discovery, disease classification, and similarity query processing on high-throughput MS/MS data of inborn errors of metabolism. *J Biomol Screen* 11:90–9.

Bennett L, Ciaffoni L, Denzer W, Hancock G, Lunn AD, Peverall R, Praun S, Ritchie GAD. (2009). A chemometric study on human breath mass spectra for biomarker identification in cystic fibrosis. *J Breath Res* 3:7–9.

Bohadana A, Michaely JP, Teculescu D, Wild P. (2008). Reproducibility of exhaled nitric oxide in smokers and non-smokers: relevance for longitudinal studies. *BMC Pulm Med* 8:4.

Buszewski B, Ulanowska A, Ligor T, Denderz N, Amann A. (2009). Analysis of exhaled breath from smokers, passive smokers and non-smokers by solid-phase microextraction gas chromatography/mass spectrometry. *Biomed Chromatogr* 23:551–6.

Cope K, Risby T, Diehl AM. (2000). Increased gastrointestinal ethanol production in obese mice: implications for fatty liver disease pathogenesis. *Gastroenterology* 119:1340–7.

Dolch M, Frey L, Hornuss C, Schmoelz M, Praun S, Villinger J, Schelling G. (2008). Molecular breath-gas analysis by online mass spectrometry in mechanically ventilated patients: a new software-based method of CO_2-controlled alveolar gas monitoring. *J Breath Res* 2:10–12.

Futrell, J.H. (1986). *Gaseous Ion Chemistry and Mass Spectrometry*. Hoboken, NJ, USA: John Wiley & Sons. p. 141.

Hornuss C, Praun S, Villinger J, Dornauer A, Moehnle P, Dolch M, Weninger E, Chouker A, Feil C, Briegel J, Thiel M, Schelling G. (2007). Real-time monitoring of propofol in expired air in humans undergoing total intravenous anesthesia. *Anesthesiology* 106:665–74.

Jones A W (1985). Excretion of low-molecular weight volatile substances in human breath: focus on endogenous ethanol. *J Anal Toxicol* 9:246–50.

Jones AW, Ostrovsky YuM, Wallin A, Midtvedt T. (1984). Lack of differences in blood and tissue concentrations of endogenous ethanol in conventional and germfree rats. *Alcohol* 1:393–6.

Jordan A, Hansel A, Holzinger R, Lindinger W. (1995). Acetonitrile and benzene in the breath of smokers and non-smokers investigated by proton transfer reaction mass spectrometry (PTR-MS). *Int J Mass Spectr Ion Process* 148:1–3.

Kharitonov SA, Yates D, Robbins RA, Logan-Sinclair R, Shinebourne EA, Barnes PJ. (1994). Increased nitric oxide in exhaled air of asthmatic patients. *Lancet* 343:133–5.

Kononenko I. (1994). *Estimating Attributes: Analysis and Extensions of Relief*. Berlin, Heidelberg, Germany: Springer. p. 171–82.

Mazzanti R, Moscarella S, Bensi G, Altavilla A, Gentilini P. (1989). Hepatic lipid peroxidation and aldehyde dehydrogenase activity in alcoholic and non alcoholic liver disease. *Alcohol Alcohol* 24:121–8.

Mazzone PJ. (2008). Analysis of volatile organic compounds in the exhaled breath for the diagnosis of lung cancer. *J Thorac Oncol* 3:774–80.

Nair S, Cope K, Risby TH, Diehl AM. (2001). Obesity and female gender increase breath ethanol concentration: potential implications for the pathogenesis of nonalcoholic steatohepatitis. *Am J Gastroenterol* 96:1200–4.

Nanji AA, Khettry U, Sadrzadeh SM. (1994). Lactobacillus feeding reduces endotoxemia and severity of experimental alcoholic liver (disease). *Proc Soc Exp Biol Med* 205:243–7.

Netzer M, Millonig G, Osl M, Pfeifer B, Praun S, Villinger J, Vogel W, Baumgartner C. (2009). A new ensemble-based algorithm for identifying breath gas marker candidates in liver disease using ion molecule reaction mass spectrometry. *Bioinformatics* 25:941–7.

Oneta CM, Lieber CS, Li J, Ruttimann S, Schmid B, Lattmann J, Rosman AS, Seitz HK. (2002). Dynamics of cytochrome P4502E1 activity in man: induction by ethanol and disappearance during withdrawal phase. *J Hepatol* 36:47–52.

Periago JF, Morente A, Villanueva M, Luna A. (1994). Correlation between concentrations of n-hexane and toluene in exhaled and environmental air in an occupationally exposed population. *J Appl Toxicol* 14:63–7.

Phillips M, Herrera J, Krishnan S, Zain M, Greenberg J, Cataneo RN. (1999). Variation in volatile organic compounds in the breath of normal humans. *J Chromatogr B Biomed Sci Appl* 729:75–88.

Quinlan RJ. (1993). Programs for machine learning. In: Quinlan RJ. *Morgan Kaufmann Series in Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers. C4.5.

Risby TH, Sehnert SS. (1999). Clinical application of breath biomarkers of oxidative stress status. *Free Radic Biol Med* 27:1182–92.

Sehnert SS, Jiang L, Burdick JF, Risby TH. (2002). Breath biomarkers for detection of human liver diseases: preliminary study. *Biomarkers* 7:174–87.

Seitz HK, Stickel F. (2007). Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat Rev Cancer* 7:599–612.

Solga SF, Alkhuraishe A, Cope K, Tabesh A, Clark JM, Torbenson M, Schwartz S, Magnuson T, Diehl AM, Risby TH. (2006). Breath biomarkers and non-alcoholic fatty liver disease: preliminary observations. *Biomarkers* 11:174–83.

Spinucci G, Guidetti M, Lanzoni E, Pironi L. (2006). Endogenous ethanol production in a patient with chronic intestinal pseudo-obstruction and small intestinal bacterial overgrowth. *Eur J Gastroenterol Hepatol* 18:799–802.

Stone BG, Besse TJ, Duane WC, Evans CD, DeMaster EG. (1993). Effect of regulating cholesterol biosynthesis on breath isoprene excretion in men. *Lipids* 28:705–8.

Takahashi T, Lasker JM, Rosman AS, Lieber CS. (1993). Induction of cytochrome P-4502E1 in the human liver by ethanol is caused by a corresponding increase in encoding messenger RNA. *Hepatology* 17:236–45.

Takalo R, Piirila P, Sovijarvi AR. (2008). Repeatability of successive measurements with a portable nitric oxide analyser in patients with suggested or diagnosed asthma. *Scand J Clin Lab Invest* 68:830–2.

Tangerman A, Meuwese-Arends MT, van Tongeren JH. (1983). A new sensitive assay for measuring volatile sulphur compounds in human breath by Tenax trapping and gas chromatography and its application in liver cirrhosis. *Clin Chim Acta* 130:103–10.

Tegtmeyer U, Weiss HP, Schlögl R. (1993). Gas analysis by IMR-MS: a comparison to conventional mass spectrometry. *Fresenius J Anal Chem* 347:263–8.

Van den Velde S, Nevens F, Van Hee P, van Steenberghe D, Quirynen M. (2008). GC-MS analysis of breath odor compounds in liver patients. *J Chromatogr B Analyt Technol Biomed Life Sci* 875:344–8.

Wallace L, Pellizzari E, Hartwell TD, Perritt R, Ziegenfus R. (1987). Exposures to benzene and other volatile compounds from active and passive smoking. *Arch Environ Health* 42:272–9.

Wang Y, Millonig G, Nair J, Patsenker E, Stickel F, Mueller S, Bartsch H, Seitz HK. (2009). Ethanol-induced cytochrome P4502E1 causes carcinogenic etheno-DNA lesions in alcoholic liver disease. *Hepatology* 50: 453–61.

Weltman MD, Farrell GC, Hall P, Ingelman-Sundberg M, Liddle C. (1998). Hepatic cytochrome P450 2E1 is increased in patients with nonalcoholic steatohepatitis. *Hepatology* 27:128–33.